DOCUMENT RESUME

ED 188 627 IR 008 496

AUTHOR Baechtold, Marguerite
TITLE Early Childhood Library Programming: Measurement and Evaluation. Final Report.
INSTITUTION Western Michigan Univ., Kalamazoo. School of Librarianship.
SPONS AGENCY Office of Education (DHEW), Washington, D.C.
PUB DATE Aug 79
GRANT G007801816
NOTE 94p.

EDRS PRICE MF01/PC04 Plus Postage.
DESCRIPTORS Audiovisual Aids; Conferences; Demonstration Programs; *Evaluation Methods; Librarians; *Library Services; Media Specialists; *Program Evaluation; Public Libraries; School Libraries; *Young Children

ABSTRACT

This research and demonstration project was undertaken to study styles and techniques of evaluation which are applicable to early childhood programming in school and public libraries, and to identify during a working conference some methods of measuring outcomes in order to improve and justify library programs for young children. The study report discusses a literature search in librarianship and the related fields of educational sociology, social work, and behavioral psychology; pre-planning for the working conference; the recruitment of participants; proceedings of the conference; and the conference follow-up. Examples of materials used in conference presentations, simulation devices, an analysis of the project, conclusions, a summary, and outcomes are provided, and appendices include a fact sheet, a revised list of staff for the conference, announcements, application and confirmation forms, a list of participants, and the agenda for the working conference. (FM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

SCOPE OF INTEREST NOTICE
The ERIC Facility has assigned this document for processing to: IR PS

In our judgement, this document is also of interest to the clearinghouses noted to the right. Indexing should reflect their special points of view.

ED188627

Final Report

EARLY CHILDHOOD LIBRARY PROGRAMMING: MEASUREMENT AND EVALUATION

OE Grant Number G007801816

August 29, 1978 - August 28, 1979
(No-cost extension to October 31, 1979)

The activity which is the subject of this report was supported in whole or in part by the U.S. Office of Education, Department of Health, Education, and Welfare. However, the opinions expressed herein do not necessarily reflect the position or policy of the U.S. Office of Education, and no official endorsement by the U.S. Office of Education is inferred.

Submitted by

Marguerite Baechtold, Project Director

School of Librarianship
Western Michigan University
Kalamazoo, Michigan

IR008496

TABLE OF CONTENTS

PURPOSE AND PROCEDURE 1

Literature Search 1

Pre-Planning for Working Conference 1

Recruitment of Participants 2

Working Conference on Measurement and Evaluation, April 17-18, 1979 2

Conference Follow-Up 5

ANALYSIS AND CONCLUSIONS 6

SUMMARY AND OUTCOMES 7

APPENDIX A. "Excerpts from the Authorities: a Fact Sheet...". . . . 9

APPENDIX B. Revised list of Staff for Working Conference 19

APPENDIX C. Announcements, Application and Confirmation Forms, and List of Participants 20

APPENDIX D. Agenda for Working Conference 26

MAR 2 1 1980

NARRATIVE REPORT:
EARLY CHILDHOOD LIBRARY PROGRAMMING
MEASUREMENT AND EVALUATION

PURPOSE AND PROCEDURE

The purpose of this research and demonstration project was: 1) to study styles and techniques of evaluation which are applicable to early childhood programming in school and public libraries and 2) to identify during a working conference some methods of measuring outcomes in order to improve and justify library programs for young children.

Literature Search - An exhaustive search in the literature of librarianship and the related disciplines of education sociology, social work, and behavioral psychology was undertaken. This effort was summarized in a document entitled "Excerpts from the Authorities: a Fact Sheet..." The "Fact Sheet" served two useful purposes: 1) to synthesize for the investigator the current state of the art re: both process and procedure of evaluation, and 2) to provide for the participants in the projected working conference a useful distillation of current thinking about measures and process of evaluation. The arrangement of the document, therefore, paralleled the contemplated format for the conference by clustering information on the following topics:

Definitions Relating to Evaluation
Characteristics of Evaluation
Why Evaluate?
What is Being Evaluated?
How Do We Evaluate?
Users of Evaluation
Words to the Wise-Advice to Evaluators

A copy of the full text of the "Fact Sheet" is included with this report in Appendix A.

Pre-Planning for Working Conference -

Planning for the conference began initially by means of telephone calls and/or personal visits to the consultants identified in the proposal. During these preliminary conversations, roles were defined, some decisions about scope and content were made and a tentative agenda was delineated.

Early in the planning, it became apparent that one of the originally identified consultants[1] would not be free to participate because of a conflicting research commitment. Illness of another of the consultants,[2] plus crippling weather necessitated a change in the projected time schedule for the conference. A strenuous effort was made to retain the particular expertise of the originally selected consultants. Competent replacements were found by drafting several persons whose combined strengths matched the requirements of the project. A revised roster of staff for the Working Conference is included with this report in Appendix B.

---

[1]Dr. Uldis Smidchens, see Proposal, p. 10.

[2]Dr. Ruth Ekstrom, see Proposal, p. 9.

Recuritment of Participants -

The participants were recruited first from among those who attended the instructional phase of the EARLY CHILDHOOD LIBRARY PROGRAMMING: IDENTIFICATION OF COMPETENCIES NEEDED project during the summer of 1978. Preliminary contacts were made by telephone with a follow-up announcement and application form. Next, general announcement flyers were distributed to persons who had participated in the survey phase of the 1977-78 project. Some difficulty was encountered in recruiting participants for the Working Conference because of limitations on released time imposed by school and public library employers. Copies of announcements, application, and confirmation forms plus a list of the participants are included with this report in Appendix C.

Working Conference on Measurement and Evaluation, April 17-18, 1979 -

In finalizing the agenda, the Consultants and Resource Persons concurred with the Project Director's goal of maximum flexibility and opportunity for interaction within the prescribed scope of the conference topic. Each of the following sections began with a presentation or disucssion, followed by a work session.

WHY EVALUATE?

Why do we need to, why do we want to evaluate?

WHAT IS TO BE EVALUATED?

What do we need to know?

HOW TO EVALUATE?

Techniques, tools, method

DISSEMINATION?

Who will need, who will be using the evaluation?

PRODUCT?

Formats, styles

Varied approaches were employed during the presentations. These included: slide/cassette shows, concept development, free association/brainstorming, and simulation activities. The work sessions were highly productive because at these times participants were able to relate the preliminary presentations to their own situations. The emphasis was on helping them to draw from the methodologies of evaluation in the related fields - sociology, education, psychology, public administration management and marketing. The ratio of staff to participants permitted a high degree of availability and interaction. The built-in flexibility provided opportunity for the consultants and resource persons to confer frequently. This permitted re-direction and change of pace as needs surfaced among the participants. The pages which follow are examples of the materials used in the presentations and simulation activities.

INPUT	PROCESS	OUTPUT
List all the resources which are necessary for your program to function properly. (These include money, people, facilities, supplies, preconditions, etc.)	Describe briefly the activities which your program is engaged in, or those in which your program intends to engage in to achieve your program's goals.	List the major goals your program will meet. These may be in terms of knowledge, skills attitudes, conditions, etc.

SLIDE SHOW 1

DISCREPANCY EVALUATION MODEL

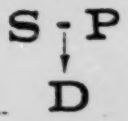
The Discrepancy Evaluation Model provides information for improving and assessing educational programs. It involves the comparison of some performance with a standard.

P = PERFORMANCE
S = STANDARD

Any difference found to exist is called a discrepancy. Discrepancies may be positive, where performance exceeds the standard, or negative, where it falls short.

S - P
↓
D (DISCREPANCY)

DISCREPANCY EVALUATION MODEL



LET'S LOOK AT AN EXAMPLE . . .

Colonel Standard realizes that there are three types of action he can take:

He can control performance.

He can redesign the standard.

Or, he can terminate the program.

SUMMARY

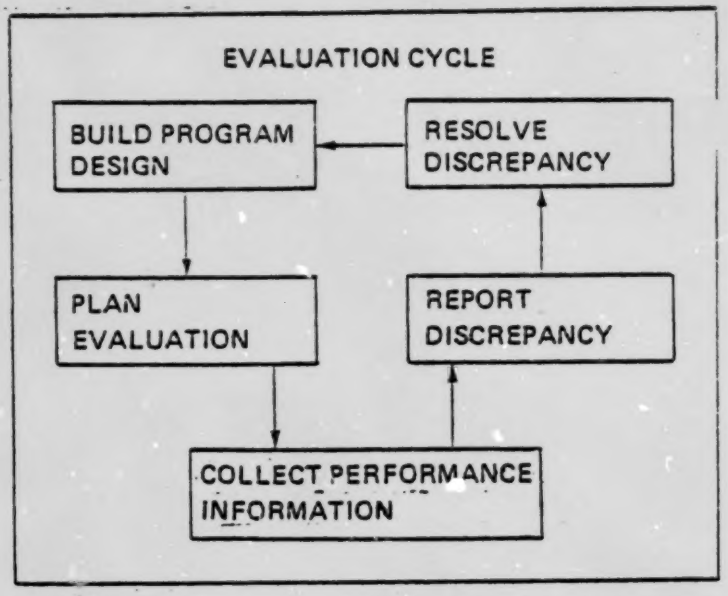
1. CONTROL – PERFORMANCE
2. REDESIGN – STANDARD
3. TERMINATION – PROGRAM

DISCREPANCY EVALUATION CYCLE

Usually, evaluation activities are applied in cycle, each step building on the next.

Evaluation is impossible without an explicit statement of program intent (standard). Standards are contained in the program design.

An educational program can be viewed as a system where processes use inputs to produce outputs. Inputs are the raw materials; processes are the program's activities; and outputs are the program's goals. Program designs are built through input, process and output analysis.



Then, evaluation planning can begin. Areas of concern are isolated and evaluation questions formulated.

The systematic collection of performance information takes place once standards have been clarified and areas of concern identified.

Once collected, discrepancy data is reported to program decision-makers, who resolve discrepancies through control, redesign or termination.

These changes may, in turn, effect the program design; therefore, the evaluation cycle may begin anew.

S reflects INTENT of
- program planners
- staff
- funding agents
- others

S is an agreement on intended
- inputs
- processes
- outputs

HUMAN RESOURCES
- FACULTY
- STAFF
- STUDENTS

MATERIAL RESOURCES

COUNSELING
COURSEWORK
PRACTICA
SEMINARS

INDIVIDUAL CHANGES
- KNOWLEDGE
- BEHAVIORS
- ATTITUDES
- AWARENESS

INSTITUTIONAL CHANGES

SHAPEMUP READING PROGRAM

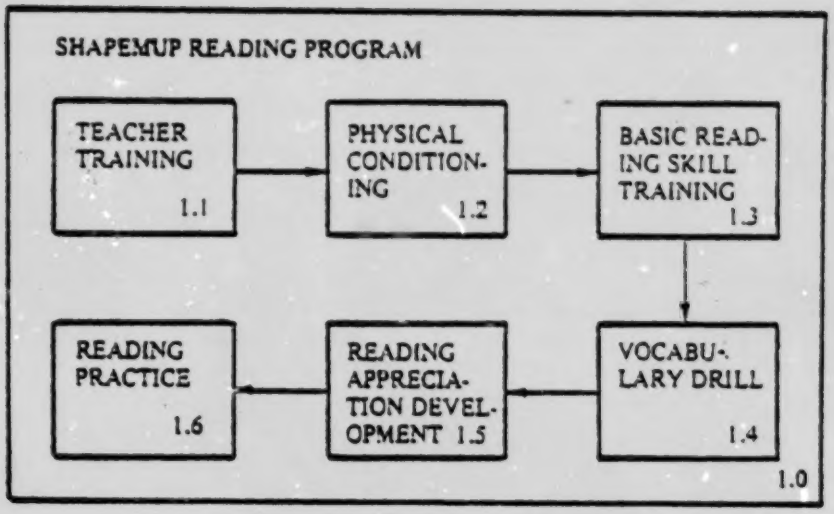
- 20 -

ANSWER KEY

INPUTS	PROCESSES	OUTPUTS
A gym Prof. of Calisthenical Communication Basic Reading Skill Teachers of sturdy physical stature	*#1. Teacher Training Component* Professor trains teachers by conducting exercise classes which illustrate proper physical conditioning techniques, and simultaneously lecturing on Calisthenical Communication theory.	Basic Reading Skill Teachers who are knowledgeable about physical conditioning techniques and Calisthenical Communication theory
A gym Basic Reading Skill Teachers who are knowledgeable about physical conditioning and Calisthenical Communication theory. Students with low Exacto scores	*#2. Physical Conditioning Component* Basic Reading Skill Teachers conduct exercise classes for students.	Students with proper physical conditioning
Instructional materials Basic Reading Skill Teachers A classroom Students with proper physical conditioning	*#3. Basic Reading Skill Training Component* Basic Reading Skill Teachers instruct students sequentially in the skills of letter recognition, phonics, and word recognition.	Students with basic reading skills
classroom Lists of 100 4-letter vocabulary words Students with basic reading skills Advanced Reading Skill Teachers	*#4. Vocabulary Drill Component* Advanced Reading Skill Teachers distribute vocabulary lists to students, review definitions, and drill students orally until they give correct responses to all words.	Students with a vocabulary of at least 100 4-letter words
Advanced Reading Skill Teachers Students with a vocabulary of at least 100 4-letter words Drama students who love to read and have expressive voices Interesting books with a high frequency of 4-letter words (to be read aloud) An air-conditioned library within walking distance	*#5. Reading Appreciation Development Component* Advanced Reading Skill Teachers take students to the library where, after listening to drama students read interesting books aloud, they choose a library book to take back to class.	Students with a book and the desire to read
Advanced Reading Skill Teachers A hot classroom Students with a book and the desire to read	*#6. Reading Practice Component* Each student reads at least 5 library books and is quizzed by an Advanced Reading Skill Teacher on its content before returning to the library for another book; after 5 books have been read the Exacto Reading Proficiency test is administered.	Students with substantially improved reading proficiency (as documented by a statistically significant increase in student Exacto scores

(Network on reverse side)

PROGRAM DESIGN NETWORK

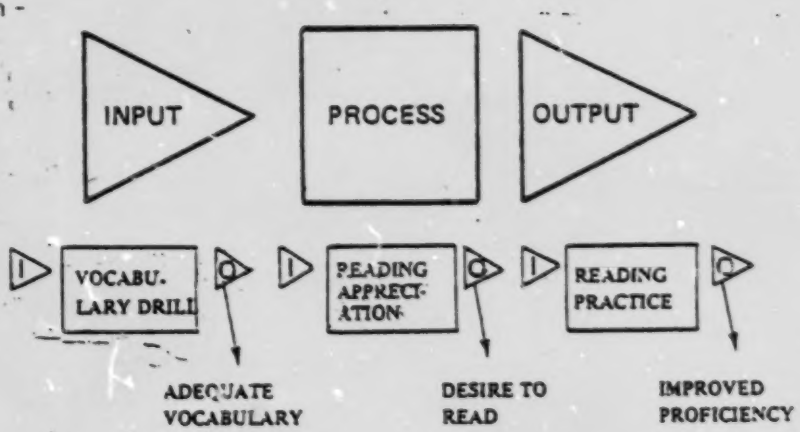


SLIDE SHOW 2

AN EDUCATIONAL PROGRAM AS A SYSTEM

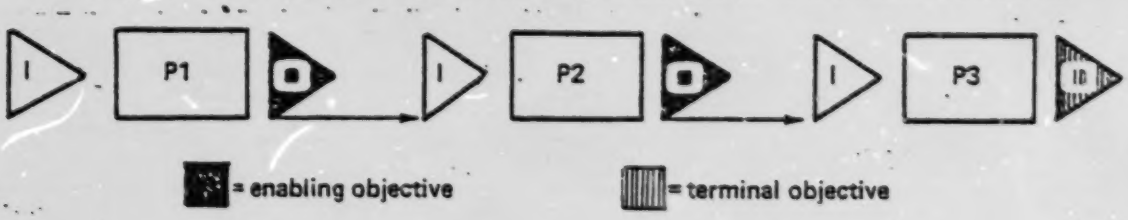
Program designs are built through systems analysis. An educational program can be seen as a single system, where processes use inputs to produce outputs.

The Shapemup Program has been broken down into interdependent sub-systems. The output of each unit enables the next unit to occur, ultimately producing the program's terminal objective.



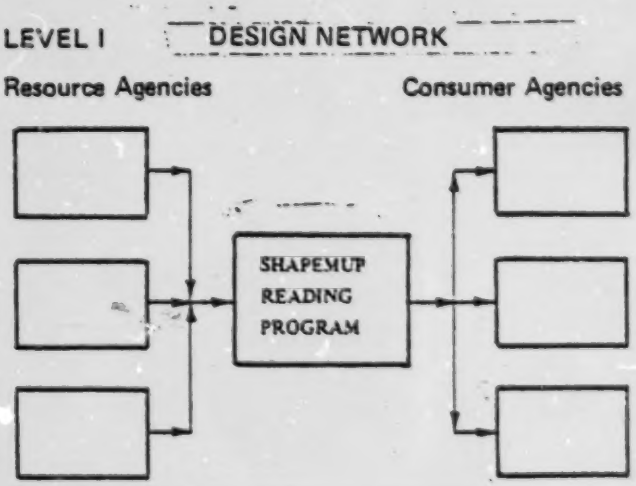
An enabling objective is the output of a component process which is needed as another component's input. An enabling objective becomes an input for the next component, and that component's output becomes an input for the next component, etc., until the program has achieved it's ultimate goal.

Looking back at the Shapemup Program, you should be able to name two enabling objectives,

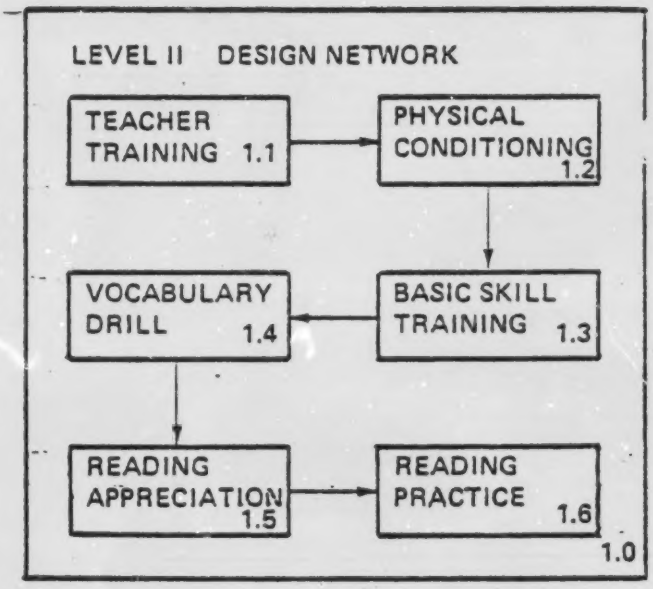


and the process that each enabling objective is dependent upon.

The most general level of a program design describes the program's connection with its environment—a relationship where all resources come out of the environment and all terminal objectives return to it.



Level II diagrams the program's major components. The arrows show that the output of one component is used as the input for the next.



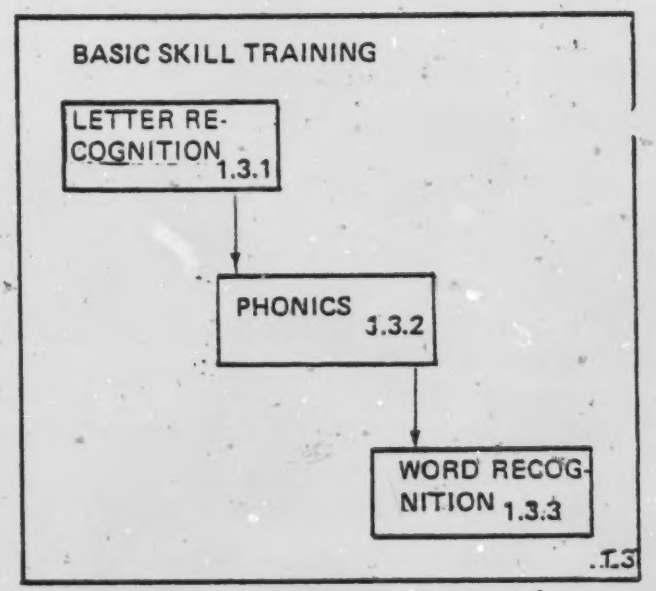
For the Shapemup Program to work, six components are needed:

1. Teacher Training—teachers trained in calisthenic teaching skills;
2. Physical Conditioning—students exercise in gym;
3. Basic Skill Training—three major skills taught;
4. Vocabulary Drill—fluency with four-letter words expanded;
5. Reading Appreciation—desire to read is developed in students;
6. Reading Practice—students with desire to read and a book practice.

When students exit from the Reading Practice component, and therefore from the program, they should have increased their reading proficiency—which is the program's terminal objective.

Each component in the program can be further analyzed into its internal operations, or sub-components.

LEVEL III DESIGN NETWORK



Consider component 1.3 Basic Skill Training. To complete this component, students are trained in three related skills. They are instructed in Letter Recognition, and when they have mastered that skill, they learn how to use their letter recognition skills in the Phonics sub-component. Then, they proceed to the next activity, Word Recognition. By completing that activity, they have completed Basic Skill Training and are ready to move on to the next major component, Vocabulary Drill where their newly acquired skills will be refined and expanded.

Evaluation planning usually takes place at this finest level of detail where inputs, processes and outputs can be isolated and examined.

The inputs might be examined to see if the necessary resources have been allocated for the activity. Or, process evaluation might occur, where the evaluator asks whether the intended activities actually were conducted. And, output evaluation could be done; in this case using tests to measure whether students have achieved the expected level of reading skill.

SHAPEMUP READING PROGRAM DESIGN SELF-EVALUATION

Directions: This form provides, in the left-hand column, a series of program design term definitions or concepts. Opposite each definition or concept are related questions which refer to the Shapemup program. Read each definition or concept, then as a group, complete the answers to the questions as best you can, referring to your Shapemup design. Your group should only complete one form.

DEFINITION/CONCEPT	QUESTIONS

Program design involves program description through successively more detailed levels of analysis. At any level the program design document consists of two parts: (1) a network of boxes, each box representing a program function; and (2) a set of input, process, and output descriptions (IPO's), one IPO for each box in the network.

The first , Level I, establishes the boundaries of the program in its environment and depicts the inputs to the program from the environment, and the outputs of the program returning to the environment.

The program at Level I is envisioned as a single function.

Level II analysis redefines this single function as a set of smaller functions (components) and describes the IPO for each of the components.

Level III analysis redefines each component as a set of smaller functions (sub-components) and describes the IPO for each sub-component.

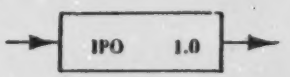
Level IV analysis redefines sub-components into sets of smaller functions and describes IPO's for the smaller functions.

At any level of analysis beyond Level I, you will usually discover that the program functions you have defined are interdependent. Arrows are placed on the network to show these dependencies. However, what is meant by a dependency?

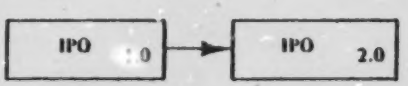
1. The design work you completed on Shapemup Reading represents:
   a. Level I analysis
   b. Level II analysis
   c. Level III analysis
   d. Level IV analysis

2. Which of the following graphic abstractions represents the analysis you performed on the Shapemup program?

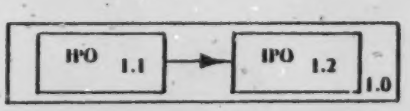
a.



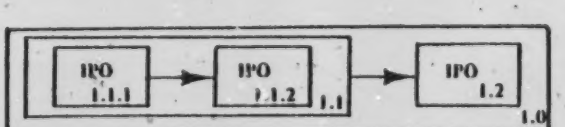
b.



c.



d.



3. Re-examine your solution to the puzzle. By our definition one statement below is the basis for placing an arrow (dependency). Which is it?
   a. Two components share the same inputs
   b. An output of one component serves as an input to another component
   c. Two processes share the same dependent variables
   d. Students flow from one component to another

24 - 15

DEFINITION/CONCEPT

An **enabling objective** *is defined as a "planned output of a program component which is needed as an input by another program component."*

A **terminal objective** *is defined as "a planned output of the program* **not** *needed as an input to any other program component, but which is desired simply as a goal."*

A **precondition** *is defined as "a qualification or characteristic of an input which is deemed as necessary to exist for the process to occur appropriately."*

QUESTIONS

4. Which of the following examples constitutes a *dependency* between components?
   a. A gym is necessary for "Physical Conditioning" and "Teacher Training"
   b. Calisthenical Communications theory underlies "Teacher Training" and "Physical Conditioning"
   c. "Students with basic reading skills" are both an *input* and *output*
   d. Any of the above examples constitute a dependency.

5. Which of the following are *enabling objectives* (circle each that applies)?
   a. Students with basic reading skills
   b. Lists of 100 4-letter vocabulary words
   c. Advanced Reading Skill Teachers
   d. Students with a vocabulary of at least 100 4-letter words
   e. Students with a book and a desire to read
   f. Teachers of sturdy physical stature

6. "Students with basic reading skills" is best described as:
   a. An input
   b. An output
   c. An enabling objective
   d. b and c above
   e. a, b and c above

7. The terminal objective(s) of this program is/are:
   a. Student reading proficiency
   b. Student reading proficiency and physical conditioning
   c. Student desire to read
   d. b and c above

8. The precondition(s) for entering students is/are:
   a. A certain minimal level of physical fitness
   b. Not specified
   c. A low Exacto score
   d. a and c above

DEFINITION/CONCEPT

Program evaluation can be applied to any components of a program and to the preconditions, enabling objectives, other inputs, processes or terminal objectives of the program. Where evaluation is applied depends, at least in part, on the purposes of the evaluation.

QUESTIONS

Here are some evaluation questions; questions which could guide evaluation activity in determining certain program effects.

Evaluation Questions

A. To what extent have Basic Reading Skill Teachers acquired the basics of Calisthenical Communications theory?

B. What pre-post shift in Exacto scores has occured?

C. What are the entering students' Exacto scores?

D. Are students listening while being read to in the library?

E. What is the correlation of vocabulary acquisition with reading rate?

Refer to the Evaluation Questions in the box above to respond to the following questions (9-12

9. To determine the program's overall effect, which question is most appropriate?

a. Question A
c. Question C
b. Question B
d. Question D
e. Question E

10. To assess the outcome of student selection, one might pursue:

a. Question A
c. Question C
b. Question B
d. Question D
e. Question E

11. What is the likelihood for success in the Physical Conditioning component?

a. Question A
c. Question C
b. Question B
d. Question D
e. Question E

12. Which question is aimed at process evaluation?

a. Question A
c. Question C
b. Question B
d. Question D
e. Question E

DISCRIMINATING EVALUATION & DESIGN QUESTIONS

Your ability to design an effective evaluation depends, in large part, upon your ability to ask good Evaluation Questions about your program design.

There are essentially two (2) types of questions one can ask about an educational program that may become confused. These are:

1. Evaluation Questions (E) —these ask *how well is the planned program working?*
2. Design Questions (D) — these ask *how should a program be planned to work?*

Here is a list of questions about the Shapemup Reading Program you just worked with. Read each, then decide whether it is an Evaluation Question (E) or a Design Question (D), and code each accordingly.

Code	Question
_____	1. Are teachers using Calisthenical Communication theory in the conditioning of students?
_____	2. How well did students perform in Reading Practice?
_____	3. Who should decide on when students will complete vocabulary drill?
_____	4. Did Exacto scores increase?
_____	5. Should Advanced Reading Skill Teachers receive instruction in Calisthenical Communication Theory?
_____	6. How low an Exacto score should a student have to be accepted into the program?
_____	7. Did the physical conditioning component produce the desired results?
_____	8. Did reading proficiency develop as a result of reading practice?
_____	9. Would increased practice raise Exacto scores higher?
_____	10. Would the Basic Skills component benefit from a more individualized approach?

Some Kinds of Measurement Procedures

Kind of Measurement Procedure	How it Works	What it Records or Categorizes
1. Observe and record behaviors of others	Observer records the behaviors of person(s) in a particular setting or time interval; behaviors are categorized or counted. Setting may be "natural" or simulated. Judgments of quality are not made.	Behaviors, actions, or events
2. Record own behaviors	Respondent maintains a record of events or behaviors involving himself, indicating nature of and/or time spent in activities as they transpire.	
3. Conduct a survey	Respondent records or categorizes events, circumstances, environmental variables, etc., as they apply to himself or others. Judgments of same are not made.	
4. Administer objective tests	Respondent selects or generates responses to given questions intended to assess knowledge, understanding, cognitive variables. Usually self-administered.	What person(s) think, know or feel
5. Collect self-ratings	Respondent records or categorizes own opinions, attitudes values or judgments about himself.	
6. Collect ratings/judgments about others	Respondent records or categorizes judgments about quality or characteristics of some event or person. NOTE: may be reflective or based on immediate observed experience.	
7. Analyze products: a. a work sample b. a product derived from a simulation	Respondent analyzes some document or product in order to determine the extent to which it contains certain elements or meets certain criteria.	Characteristics of tangible objects
8. Analyze existing records or archives	Usage reports, receipts, etc., are analyzed, counted, documented or aggregated.	
9. Produce an inventory	Respondent counts, measures or categorizes certain tangible objects and records results.	

CRITERIA for Selecting Measurement Procedures

A specific measurement procedure should indicate **what** is to be done, **who** will be doing it, and **when** or **how** it will be done.

There are three main criteria for judging the quality of a measurement procedure. The procedure should: **1)** be practical in its implementation; **2)** be ethical; and **3)** elicit credible data.

1. Practicality

When selecting or developing measurement procedures, time commitments, development and implementation costs, and political concerns should all be considered.

Time commitments
How much time will be required to carry out the measurement procedure?

Will instruments need to be developed?

Will staff need to be recruited and/or trained?
How much time will be required to collect, aggregate, code, analyze, and store the data?

Development and implementation costs:
How much will it cost to develop the instrument?

How much will it cost to train staff and administer the instrument?

What scoring mechanism will be used? Hand or machine?

Political Concerns:
Who will implement the measurement procedure? Will it inconvenience them to do so?

Who will be responsible for developing the instrument?

Are there individuals or groups that might be opposed to this procedure?

What possible positive or negative effects will the measurement procedure have on the respondents? the program? the staff?

2. Ethics

This criterion is related to the legal and moral implications of the measurement procedure.

Human rights
Does the measurement procedure violate any personal rights of privacy, equal protection, etc.?

Legality
Does the measurement procedure violate any law or regulation?

Confidentiality/integrity
Will the information collected by the measurement procedure be kept as confidential and anonymous as necessary to protect human rights?

3. Credibility

Evaluation data must be **credible.** They must convincingly represent whatever aspect of the program's performance they purport to assess. Factors affecting credibility are
Validity
Reliability
Objectivity
Reactivity
Bias

Note
Validity and Reliability: while these are terms more usually associated with instruments themselves, careful design of the entire measurement procedure can maximize these factors.

Validity
This relates to the extent to which data in fact represent what they purport, or are intended, to represent. Validity derives from the relationship between exactly what is measured and to what one wished to make an inference. For this reason, selection of the measurement procedure is extremely important to validity.

Is what the procedure will measure logically related to the dimensions and evaluation questions being addressed by that measurement procedure?

Is there reason to believe that differences reflected by the data collected will reflect real differences in the traits, characteristics, or skills about which information is sought?

How directly will the procedure measure the dimension?

Reliability
This pertains to the accuracy of the measurement. How accurately will it measure whatever it is measuring?

Will the procedure be adversely affected by any peculiar characteristics of a particular measurement setting?

Can the measurement procedure be implemented consistently from instance to instance?

Can respondents make required judgments or categorizations accurately?

Objectivity
This relates to the amount of subjectivity inherent in, or caused by, the measurement procedure. It related critically to the 'believability' of the resultant data.

Will respondents make required judgments or categorizations honestly?

Will respondents perceive a premium for responding in a particular way?

Reactivity
The measurement procedure should not adversely affect (change in any serious way) what it is attempting to measure.

Will the measuring instrument 'teach' certain responses?

Will observers or recorders adversely affect what is to be measured?

Will particular demands of the measurement procedure adversely affect the object of measurement?

Bias
The method by which the measurement procedure samples respondents and/or objects—of—measurement (behaviors, records, products, etc.) should be designed so as not to yield unusual or unrepresentative samples.

Do respondents 'self—select'?

Will samples be representative?

Will raw data finally available for analysis be representative of only one group or point of view?

DIMENSION NOTES

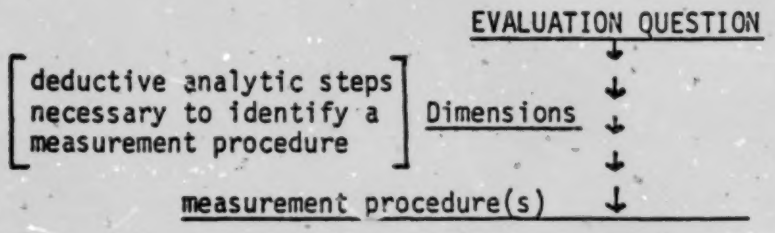
An evaluation question poses a potential comparison of program performance to a standard at a relatively general level. Because it is relatively general, it is open to various interpretations. The measurement procedure used to measure the evaluation question will greatly restrict the meaning of the evaluation question. Once a measurement procedure is selected, the meaning of the evaluation question is defined by what the measurement procedure does.

What we measure in educational programs should be specified on the basis of the INTENT of a program, not by which measurement procedure is available or possible. Too often an evaluation question is formulated and a measurement procedure chosen before thought is given to what aspects of the evaluation question should be answered.

This is why dimensions are important. By choosing relevant dimensions, a measurement procedure can be decided upon that will 1) be an appropriate measure of the program intent, and 2) retain the meaning of the evaluation question and keep it consistent with the program design.

WHAT IS A DIMENSION

A dimension is an arbitrary term we apply to that part of the measurement planning process that further defines the evaluation question.

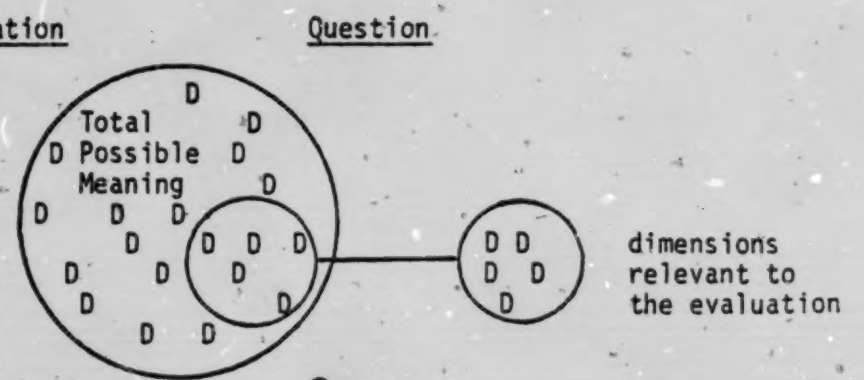


The number of steps involved in the process varies with each question. Some evaluation questions will require more analytic steps or will have more dimensions than others. The number of dimensions chosen depends on 1) how general and broad the question is to begin with, 2) the level of specificity required to address the evaluation question adequately, and 3) the personal interpretations of the evaluation question.

Go to first exercise -- Specificity

HOW DIMENSIONS ARE CHOSEN:

For any evaluation question there are many dimensions which, when combined, define its total meaning. However, for practical evaluation, only the relevant aspects of the evaluation question should be considered. Those appropriate aspects are the dimensions of the evaluation question.



12b (in)

When choosing the dimensions of an evaluation question, it may be helpful to keep in mind:

1) What is the focus of the question (i.e., About whom or what are we seeking information?)
2) Why is this question important?

By using the answers to these questions as guidelines, the relevant dimensions can more easily be chosen.

SOURCES FOR DIMENSIONS

Choosing dimensions is a process. Several sources can and should be utilized when deciding what dimensions are appropriate to your evaluation needs. One source that can help suggest dimensions is the "design referent" column of the "evaluation design" form (11 ho). This column indicates whether the evaluation has been labeled as an input, process, or output in the context of the program being evaluated.

If the evaluation question is input or output-related (or both, as in the case of an enabling objective), appropriate dimensions will depend on the focus of the question.

If the focus is on individuals, you may want to look at particular:

- knowledges
- cognitive skills
- psychomotor skills
- behaviors
- attitudes
- perceptual skills

If the focus in on an organization, you may want to look at:

- patronage patterns
- structure and roles
- goals and resources

If the focus is on a product, you may want to look at:

- intrinsic qualities
- who uses it
- how it's used
- how well it works

If the evaluation question is process-related, appropriate dimensions might be derived from:

- the use of resources
- behavior patterns of the persons involved
- particular process criteria

Another source that can help suggest dimensions is the column of the "evaluation design" form labeled "Why the question is important." This column is helpful in clarifying the focus of the questions. It also helps pinpoint critical areas from which you need to gather information to answer your evaluation questions.

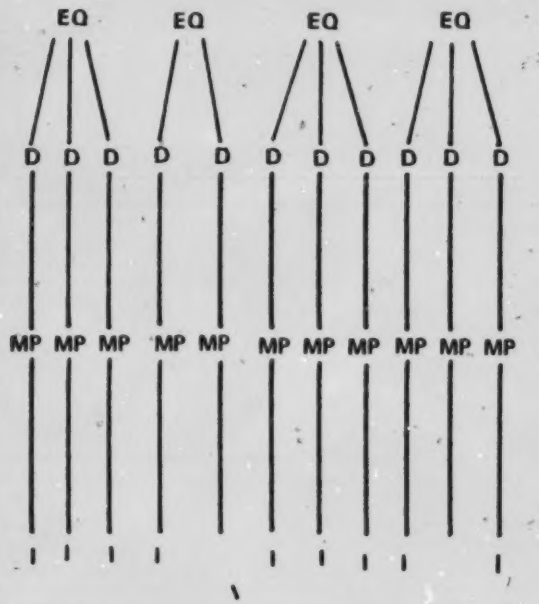
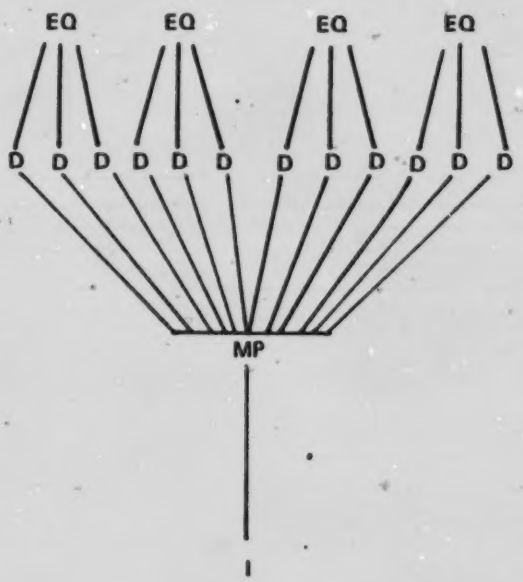
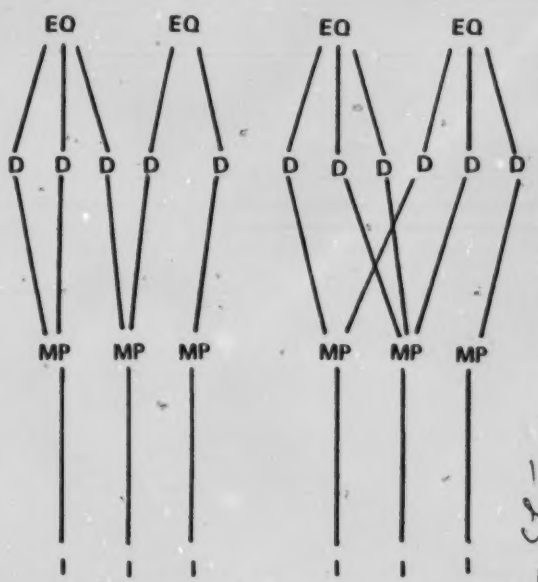
An additional source which can prove helpful to you is the list of words below. Dimensions often can suggest quantitative or qualitative characteristics and often include words like:

number	duration	nature	availability	competence
amount	existence	difficulty	relevance	knowledge
frequency	sequence	completeness	adequacy	appropriateness
size	type	sufficiency	ability	effectiveness
				utility

Evaluation Concern	Evaluation Question	Design Referent	Why the Question is Important
Successful Operation of Seminars	To what extent have students acquired intended seminar learning outcomes?	Output (2.4)	The seminar is to contribute to skill and knowledge in each of the three competency domains, depending on the nature of problems the student encounters during the practicum.
	Are the seminars being used to provide constructive guidance to the practicum?	Process (2.4)	If the seminar is not providing this, one of its major goals, it should be redesigned to ensure that students receive constructive guidance from faculty and other studen
	Is seminar attendance at the expected level?	Process (2.4)	Seminar attendance is voluntary, but the faculty had hoped for 80% attendance.
	Have seminar activities impacted on practicum learning processes?	Output (2.4)	Where a student's practicum activities have been revised and modified as a result of conclusions reached in the seminar, the modifications should enhance the practicum's impact.

DIMENSIONS → / MEASUREMENT PROCEDURES ↓	1 Is student adequately prepared to begin practicum?	2 Is the practicum site on which the student is placed adequate to meet training needs?	3 Is the student engaging in intended learning activities?	4 Has each student mastered the SSM competencies after completion of the practicum?	5 How well do Acme graduates perform on the job?	INSTRUMENTS OR DATA SOURCE
	a. Minimum knowledge of SSM techniques b. Minimum knowledge of customer types	a. Types of cars available for sale at site b. Size of inventory at site	Amount & types of: a. Maintenance activities b. Admin. activities c. Different types of sales activities d. Contacts with supervising sales person	a. Total no. of cars sold by student b. Ability to use SSM on 3 customer types c. Criterion level knowledge of SSM	a. Type of positions held b. Type of clients worked with c. On the job effectiveness	
A Administer SSM knowledge test before and after practicum to student	1a 1b			4c		HOUZON–FURST SCALE OF SSM MASTERY
B Analyze bimonthly car lot sales and inventory record		2a 2b		4a		INVENTORY AND SALES RECORD
C Student keeps a record of activities engaged in during practicum			3a 3c 3b 3d			STUDENT DAILY LOG
D Frequency analysis of faculty supervision records at the end of each month			3d			RECORD OF SUPERVISORY CONTACTS
E Observe student interact with 3 covertly role played customer types during the last two weeks of the practicum				4b		SALES BEHAVIOR OBSERVATION RECORD
F Administer questionnaire to customer after interaction with student				4b		CUSTOMER FEEDBACK FORM
G Survey employers of Acme graduates six months after the graduate begins					5a 5c 5b	ACME EMPLOYER'S RATING FORM

Measurement Plan Characteristics

–A– ultiple Data Sources	–B– Economic Measurement	–C– 'Compromise' Plan
		
Using Multiple Data Sources results in too many instruments e vertical analytical pattern), although this approach is very omprehensive.	Plan B is nice since it is very economical, but it probably is impossible. Also, there are no multiple data sources to increase the quality of information obtained	This plan is a compromise of the desirable features of A and B. It affords multiple data for each EQ and multiple usage of most instruments.

Terminology

Evaluation Question (EQ)
Identifies a particular **input, process, or output** aspect of the program's performance which will be compared to the **standard** for that performance.

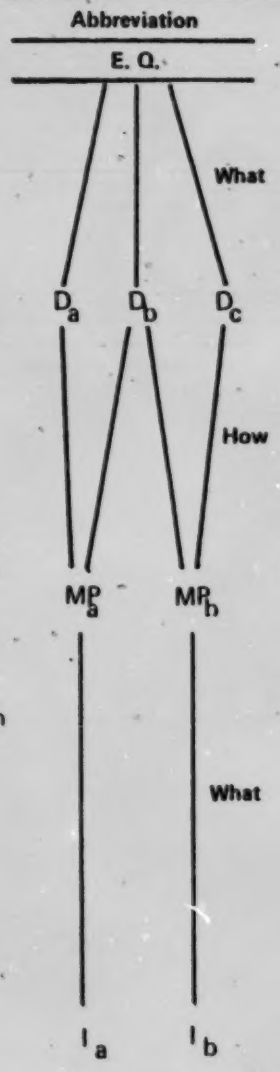
More specifically defines the S-P comparisons for each Evaluation Question

Dimension (D)
Further explicates the standard(s) implied by an Evaluation Question. Dimensions are the relevant characteristics of the Evaluation Question's object, or focus, about which one wishes to collect data in order to 'answer' the evaluation question.

Measurement Procedure (MP)
A method for collecting data about one or more dimensions. A measurement procedure specifies what kind of instrument will be used, how, by whom, and when and where in the course of the program.

Operationalizes the S-P comparisons into one or more measurement strategies

Instrument (I)
A form on which information about the dimension(s) is recorded as specified by the measurement procedure. The instrument's **items** will determine just **what** about the dimension(s) is to be recorded.



Example

Evaluation Question
'How well is the new practicum advising procedure working?'

Dimension

D_a: **Frequency** and **amount** of advising taking place

D_b: **nature** of problems being addressed by advising

D_c: **adequacy** of the advising to help with problems

Measurement Procedures

MP_a: 'Content-analyze samples of the faculty-maintained advisement records each month during the practicum

MP_b: Survey students with a questionnaire at end of each quarter of the practicum; have them describe and rate advising received.

Instruments

I_a 'Aggregate Advisement Description Record'

I_b 'Student Advisement Questionnaire'

2.0 Master's Degree Program — WHITCOMB COLLEGE SPECIAL EDUCATION DEPARTMENT
LEVEL III — INPUT-PROCESS-OUTPUT DESCRIPTION

INPUTS	PROCESS	OUTPUTS
• (40) selected students (from 5.1) • Special Education faculty (1 associate, 1 assistant professor) • general education faculty • student schedules (from 2.5) • classrooms, etc. • student guidance from learning managers relative to program of study (from 2.5)	**2.1 BASIC COURSEWORK** Students take 9 hours of foundation courses in general education and 6 hours of foundation courses in special education. All special education foundation hours and 3 of the 9 general education hours are taken during the student's first summer in the program. The remaining 6 hours in general education are taken during the following fall semester.	• students are able to relate special education to the larger context of general education • students are exposed to a wide range of education and special education theory, practice, and philosophy

INPUTS	PROCESS	OUTPUTS
• Skills and methods modules (6 topic areas X 4 areas of specialization = 24 modules) • students who have: 1. selected their area of specialization (from 2.5) 2. basic background in general and special education foundations (from 2.1) 3. guidance and evaluation from learning manager (from 2.5) 4. access to the resource center	**2.2 SKILLS & METHODS MODULES** During the fall semester, students proceed through a series of 6 one-credit hour self-instructional modules which are designed to provide them with the basic skills necessary for aiding and teaching exceptional children in the schools. The skill areas taught by the modules are the same for all students, but students must choose from among four module "sets"—one for each type of exceptionality. Each module contains components for pre-assessment, readings, skill exercises, and evaluation. Modules are generally taken sequentially. Although some flexibility in pacing is allowed, all modules must be completed by the beginning of December.	• students are aware of and able to use a wide variety of external resource and support systems for exceptional children. • students are familiar with the principles and practice of psycho-educational assessment for their area of specialization. • students are able to design and evaluate remediation strategies and individual or group compensatory lessons. • students are exposed to human relations theory relevant to effective counseling and consultation. • students are familiar with a wide variety of classroom and behavior management techniques. • students are able to render an appropriate diagnosis and prescription for a paper child.

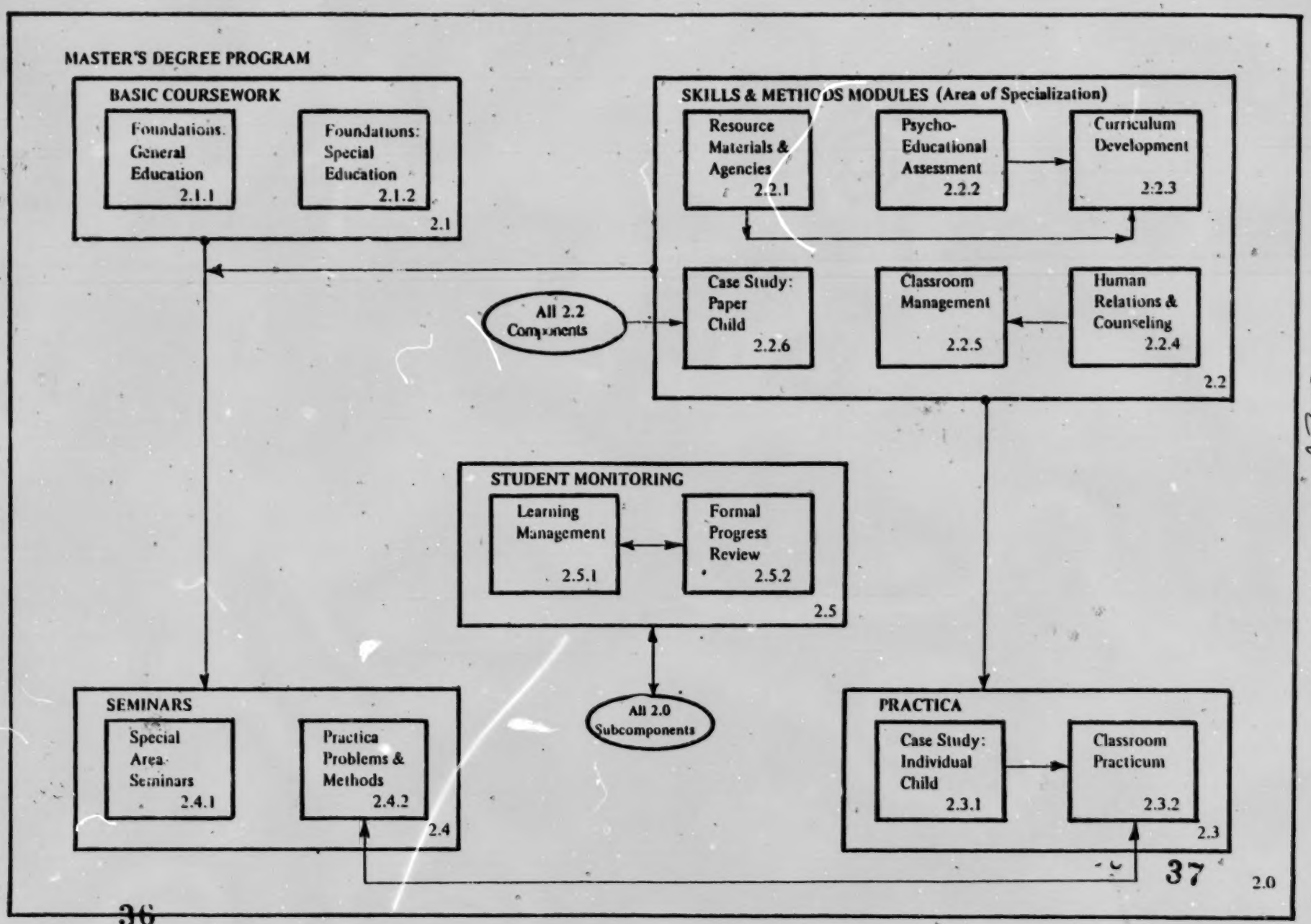
INPUTS	PROCESS	OUTPUTS
• 20 practicum sites (from 4.2) • 20 practicum site supervisors (from 4.2) • student placement in practicum sites (from 2.5) • liaison with learning manager (from 2.5) • students who have: module skills and knowledge in their area of specialization (from 2.2) • practica problems and methods seminar (from 2.4)	**2.3 PRACTICA** Students engage in two practica activities which are relevant to their chosen area of specialization. The first comes at the end of the fall semester, when students work with an individual child. The second practicum runs throughout the spring semester and involves the students on an active and responsible basis in the total classroom environment. The two practica are normally conducted at the same site and always under the supervision of a site-based professional.	• students are provided with an opportunity to apply the skills and knowledge gained in components 2.2 and 2.4 • students gain experience and competency in skill areas related to their anticipated professional roles.

INPUTS	PROCESS	OUTPUTS
• special education master's program faculty • students who have: 1. module skills and knowledge related to an area of specialization (from 2.2) 2. basic background in general and special education foundations (from 2.1) • student classroom practicum experiences (2.3) • classrooms, etc.	**2.4 SEMINARS** Students complete their coursework by taking three seminars: two oriented to issues in their area of specialization (or related areas), and a third seminar designed to complement their classroom practicum.	• students expand their knowledge base in their area of specialization • students are able to maximize the potential of their classroom practicum experience.

INPUTS	PROCESS	OUTPUTS
• students • special education master's program faculty • liaison with departmental administration • liaison with practica site personnel • student academic and practica experiences (from 2.1, 2.2, 2.3, 2.4)	**2.5 STUDENT MONITORING** Throughout the Master's program, student progress is monitored both formally and informally. The bulk of this monitoring responsibility is shouldered by the student's learning manager who guides the student's program and oversees his module instruction and practica. More formal progress review occurs at the end of the fall semester and at the completion of program coursework.	• students receive guidance relative to all aspects of their master's program. • records are kept which monitor student progress through the program

WHITCOMB COLLEGE SPECIAL EDUCATION DEPARTMENT

Level IV Network



36

37

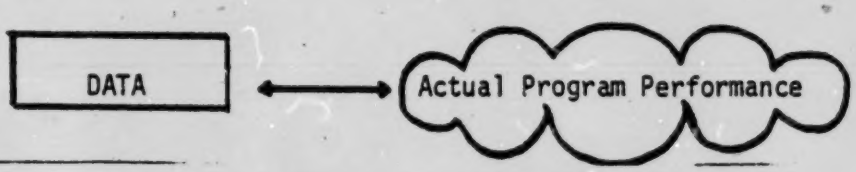
- 2v -

10b (1n)

Measurement planning follows evaluation planning. Evaluation planning involves:

- analyzing program intent and producing a program design, or standard for program operation
- generating a list of evaluation questions that ask whether particular parts of the program are in fact working as designed or having their desired impact
- determining the importance of each question, i.e., how critical it is that information be provided which will help answer the question.

The purpose of measurement is to provide information (data) about actual program performance.



Data must: 1) credibly represent program performance; and 2) be practical to collect. To ensure that data meet these criteria, careful measurement planning is needed.

Measurement planning involves:

- deciding what to measure for each evaluation question
- determining how to measure
- developing or selecting needed data collection instruments
- organizing measurement procedures into an overall data collection plan.

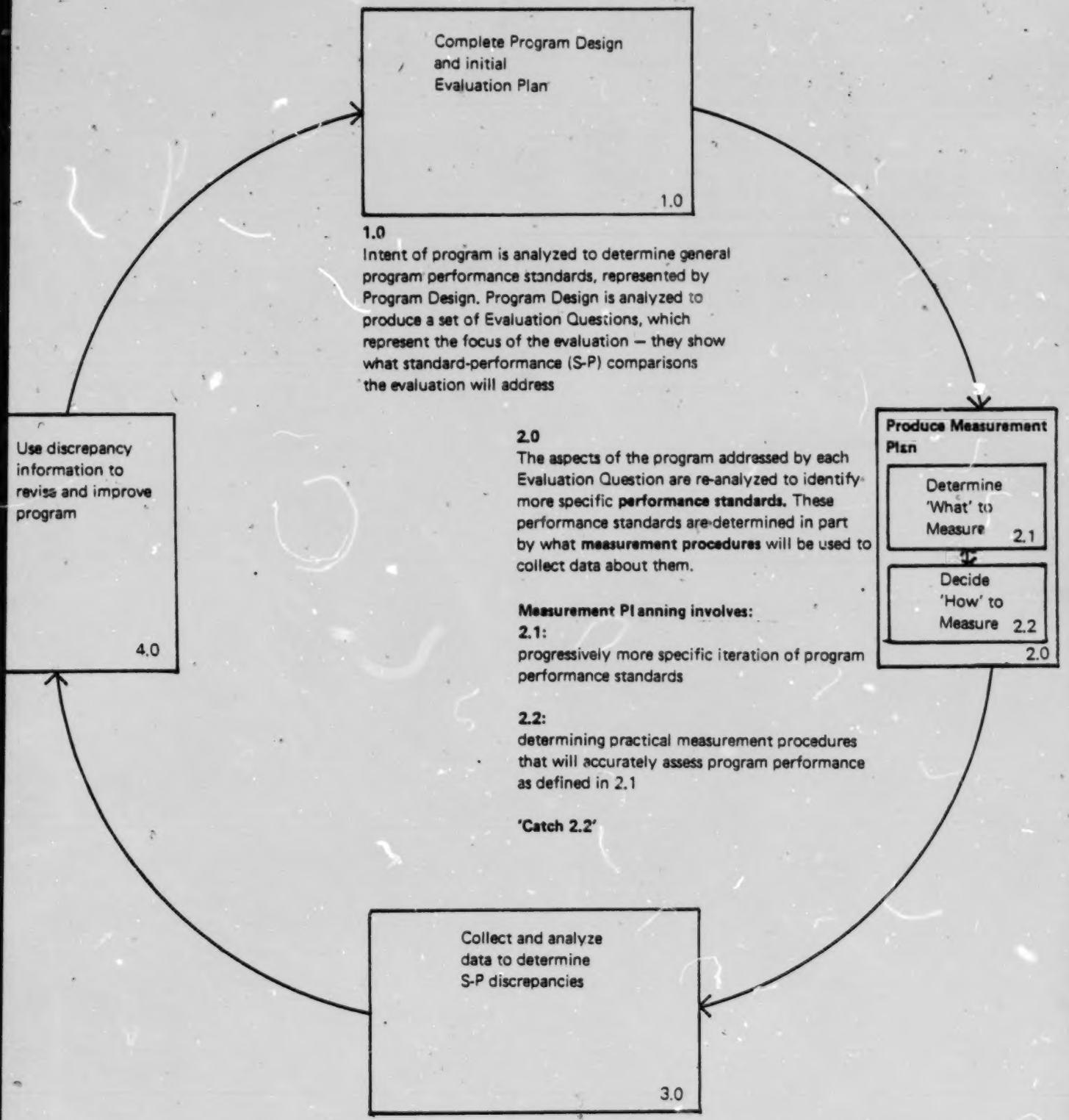
To help decide what to measure, you

1. determine the important dimensions of each critical evaluation question
2. consider what phenomena, events or characteristics of the program and its participants might be examined in order to learn about each dimension.

To help determine how to measure, you

1. Identify existing information which could address the evaluation question
2. Consider the use of existing measurement procedures and instruments
3. Identify alternative measurement procedures for unaddressed dimensions
4. Select measurement procedures which will provide accurate information about actual program performance at a reasonable cost.

Measurement Planning In The DEM Cycle


Complete Program Design and initial Evaluation Plan
1.0
Produce Measurement Plan
Determine 'What' to Measure 2.1
Decide 'How' to Measure 2.2
2.0
Collect and analyze data to determine S-P discrepancies
3.0
Use discrepancy information to revise and improve program
4.0

1.0
Intent of program is analyzed to determine general program performance standards, represented by Program Design. Program Design is analyzed to produce a set of Evaluation Questions, which represent the focus of the evaluation — they show what standard-performance (S-P) comparisons the evaluation will address

2.0
The aspects of the program addressed by each Evaluation Question are re-analyzed to identify more specific **performance standards.** These performance standards are determined in part by what **measurement procedures** will be used to collect data about them.

Measurement Planning involves:
2.1:
progressively more specific iteration of program performance standards

2.2:
determining practical measurement procedures that will accurately assess program performance as defined in 2.1

'Catch 2.2'

EVALUATION QUESTION ANALYSIS WORKSHEET

valuation Concern

valuation Question

hat is the focus of this question?

IMENSIONS:

.

.

.

.

.

.

-24-

Evaluation Concern

Evaluation Question

What is the focus of this question?

DIMENSIONS:

a.

b.

c.

d.

e.

f.

valuation Concern

valuation Question

hat is the focus of this question?

IMENSIONS:

40

.

.

.

.

Evaluation Concern

Evaluation Question

What is the focus of this question?

DIMENSIONS:

a.

b.

c.

d.

e.

f.

41

SOME TYPES OF ITEMS

Closed Items
These items consist of a stem and a set of response categories from which the respondent chooses.
Some items of this type are listed below.

1. Scales
Consists of a stem and a set of responses on a continuum.
a. rating scale — respondent checks the appropriate category.

What is your attitude toward prunes? (check one)

( )	( )	( )	( )	( )
I love them		neutral		I hate them

b. graphic scale — respondent indicates response on a line.

Indicate how you feel about prunes by placing a check in the appropriate place on the line below.

warm — cold

c. Likert scale — respondent indicates agreement or disagreement with a statement.

The prune should be America's national fruit.

( )	( )	( )	( )	( )
strongly disagree				strongly agree

2. Checklist
Consists of a set of response categories and directions for choosing response.
a. single selection — respondent selects one response among those offered.

Place a check next to the item listed below you would most like to eat.

___ a. your sock
___ b. a prune
___ c. a Bic Banana
___ d. a Detroit lemon

b. multiple selection — respondent selects all responses that apply

Which — if any — of the films listed below did you see in the past three years?

___ 'The Great Prune Robbery'
___ 'Everything You Wanted To Know About Prunes But Were Afraid To Ask'
___ 'War and Prunes'
___ 'Gone With The Prune'
___ 'The Prune That Ate Milwaukee'
___ ' Prune 1975'
___ 'The Prune Of Anne Frank'

3. Ranking — requires respondent to rank order items on a list.
Looking at the fruits listed below, rank them in terms of how much you like each. Place a '1' next to your favorite fruit, a '2' next to your second favorite fruit, and so on until all seven are ranked.

___ apple
___ blueberry
___ grapefruit
___ orange
___ plum
___ pomegranate
___ prune

4. Matched answers — respondent matches two columns of items.
For the list of occasions below, select from the list of fruits those that you would want to be served. Enter in the space in front of each occasion the letter next to the fruit of your choice.

Fruits
a. grapes
b. muskmelons
c. prunes
d. pears
e. lemons
f. raspberries

Occasions
___your birthday
___Christmas breakfast
___April Fool's Day
___IRS due date
___your triple-critical biorhythm low
___Friday afternoon

5. Paired Answers — consists of paired responses, of which the respondent chooses one.
For each pair of fruits listed below, circle the one that you prefer to eat.

a. grape or prune
b. pomegranate or prune
c. prune or green banana
d. lemon or prune
e. grape or pomegranate
f. green banana or lemon
g. grape or green banana
h. lemon or grape
i. pomegranate or green banana

Open Items

These are items for which the respondent produces the response instead of selecting one from a list provided on the instrument.

6. Short Answer — asks a question, the response to which is written on space provided.
Briefly describe in the space below your feelings about prunes.

_____

_____

7. Comments space
Often follows a closed question or set of questions, providing an opportunity for the respondent to indicate further elaboration; may also appear at the end of a questionnaire to elicit additional comments or suggestions for improvement of the instrument.

8. Fill-in the blank — requires a short answer response with the response unit specified.
Write in the space below the number of prunes you would consume each day if you were provided at no cost all the prunes you wished.

___prunes

Assume you had $100 to spend for fruit this week. List the fruits you would purchase and apportion your fruit budget across the fruits you have chosen. Choose no more than 5 fruits.

Fruits	**Amount to be spent**
_____	$ _____
_____	$ _____
_____	$ _____
_____	$ _____
_____	$ _____

SOME CONSIDERATIONS WHEN WRITING ITEMS

1. Should the items be open or closed?
The answer depends on which item type is most appropriate to elicit the desired information. Closed items are more appropriate when short responses are required and the range of responses is known. These responses can be specified by a set of response categories which are exhaustive (i. e., all possible responses are included) and mututally exclusive (i.e., response categories do not overlap). Such items are easy to code and analyze, although they restrict the information obtained to the respondent's reaction to the given response categories. When the domain of responses is not known or when detailed information is needed, open items are more appropriate. These provide more information, but they are more difficult to code and there may be some resistance on the part of the respondents to answer these items which require more time.

2. How many points should be on a rating scale?
Usually, rating scales have from three to seven points, depending on the degree to which it is possible for respondents to make the discrimination required and on the extent to which discrimination is necessary for the data analysis. That is, if it is reasonable to assume that the respondent can only discriminate among three categories on the scale, using a five- or seven-point scale is unnecessary and possibly even detrimental to the purposes of the instrument. A related issue to be considered is whether to use an even or an odd number of points on a scale which has polar opposites at its end points. Having an odd number of points has the advantage of allowing the respondent an opportunity to choose a middle category, signifying neutrality on the attribute. An advantage of having an even number of points is that the respondent can not be neutral and must choose a point on one side of the scale.

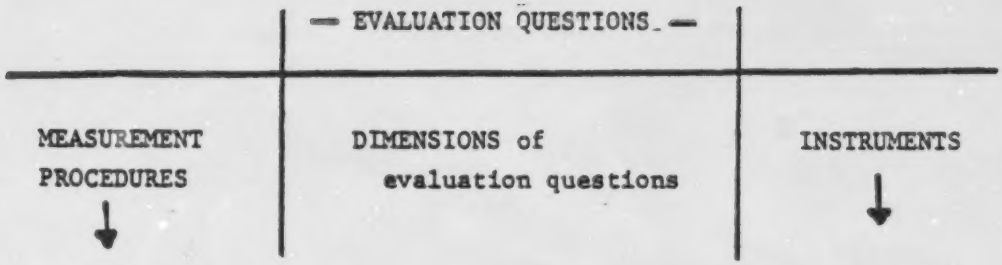
3. How completely should the anchors be specified for the response categories?
On rating scales the anchor points may be written for all points or for only a few. In the first example on page 1 the anchors for the end points and the middle category are the only ones identified, although there are a total of five response options. Some advantages in specifying all anchors is that all respondents are given the same identifiers for all categories, which is particularly important where there may be different interpretations for the categories or when respondents are not familiar with answering questionnaire items of this type. However, in some cases it is difficult to develop precise category identifiers, and instrument developers prefer to omit the name of the category , with the assumption that respondents will be less biased by having no category name than they would be by an imprecisely stated one.

4. Should numbers be written on the response categories?
For ease in coding the responses after the questionnaire is completed, it is convenient to have numerical tags for the response categories. These numbers can then readily be transferred to summary sheets for data analysis. A disadvantage of using numbers near the response category is that they may be confusing for respondents who assume these numbers are related in some way to the response categories.

5. Should statement be expressed positively or negatively?
The choice of the positive or negative phrasing of an item stem depends largely on which is the most easily understood by the respondent. For example, it may be confusing to use a Likert scale with some items in which the stem includes the word 'not' because of the confusion resulting from a double negative when a disagreement response is considered.

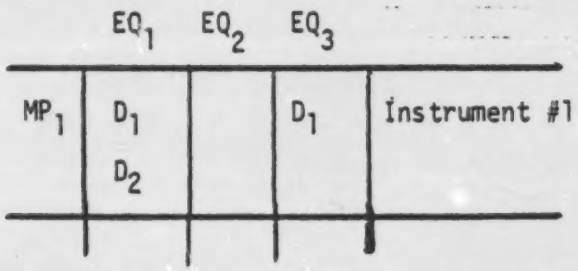
6. How many different types of items should be used on a questionnaire?
The answer to this question depends on how much of what kind of information is necessary to answer the evaluation questions. Generally, it is advisable to include several kinds of items to provide variety for the respondents. At least a few items which can be answered quickly should be in the beginning so the respondent can have a sense of accomplishment soon. More complex, or open items can be interspersed in the remainder of the questionnaire. Items which require the same response categories should be grouped. But it's good to avoid a lengthy listing of same-type items, since respondents may become impatient, and not read items carefully, or may develop a response set by selecting a favored response category and using it arbitrarily.

The Measurement Planning Matrix
- Notes -

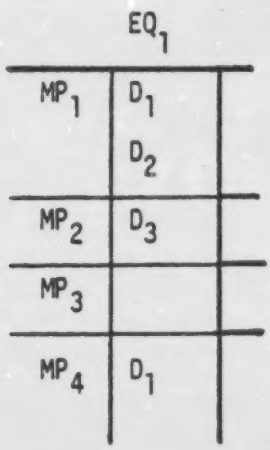


- The matrix shows how Measurement Procedures and Instruments are related to particular Dimensions of the Evaluation Questions

- Each row in the matrix shows just what dimensions are to be covered by a measurement procedure. Thus, the content for each data collection instrument can be derived from the matrix.



- Each column in the matrix shows what measurement procedures will apply to any evaluation question. The "sum" of that column constitutes an operational definition of that part of the program addressed by the evaluation question.



- Therefore, any given cell-entry shows that the dimension entered is:

  - to be measured by the procedure in the row in which it is entered

  - is a part of the evaluation question in the column in which it is entered

General purpose of the Matrix:

Used to determine adequacy of overall measurement plan - helps to ensure that evaluation questions will have a sufficient data base from multiple measurement procedures and that measurement procedures can be used to address more than one evaluation question.

Some Potential Problems a Matrix can Identify:

Signal	Potential Problem
few cell entries in a given column	an "under-measured" evaluation question
few cell entries in a given row	a "low-yield" measurement procedure
too many entries in a given column	an "over-measured" evaluation question
too many entries in a given row	an "over-loaded" measurement procedure

TERMINOLOGY

Evaluation Questions (EQ)

Identifies a particular input, process, or output aspect of the program's performance which will be compared to the standard for that performance.

Dimension (D)

Further explicates the standard(s) implied by an Evaluation Question. Dimensions are the relevant characteristics of the Evaluation Question's object, or focus, about which one wishes to collect data in order to "answer" the evaluation question.

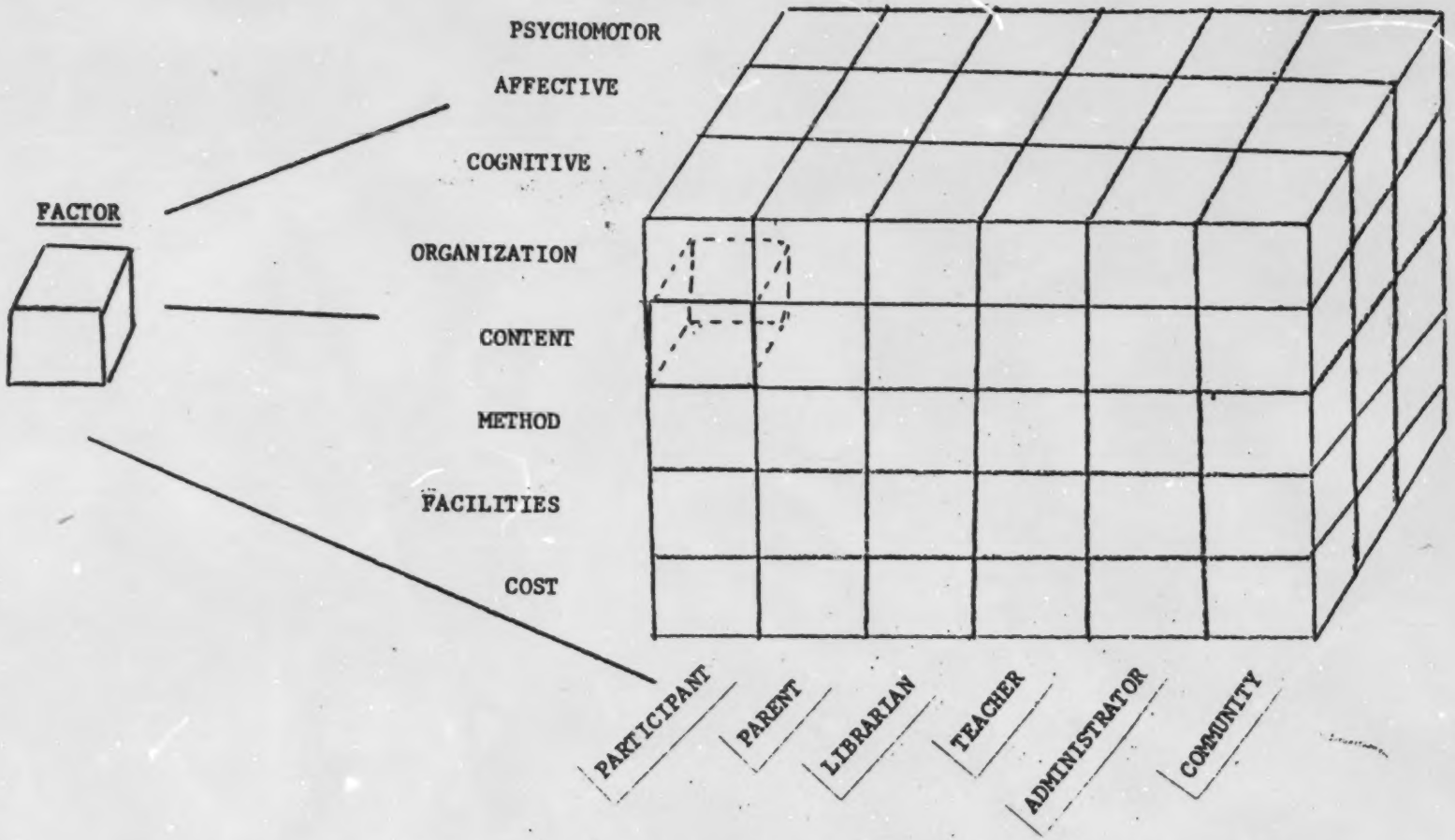
Measurement Procedure (MP)

A method for collecting data about one or more dimensions. A measurement procedure specifies what kind of instrument will be used, how, by whom, and when and where in the course of the program.

Instrument

A form on which information about the dimension(s) is recorded as specified by the measurement procedure. The instrument's items will determine just what about the dimension(s) is to be recorded.

A STRUCTURE OF VARIABLES ALONG THREE DIMENSIONS



Adapted from Booklet for Developing Evaluative Skills, Armstrong, et al, Educational Innovators Press, Tucson, Arizona, 1971.

KALAMAZOO VALLEY COMMUNITY COLLEGE

SP Evaluation Plan - 1975 - 76
oject Manager: Nancy Woods
epared: September 4, 1975

Figure 2

Reference:
MISSION:
GOAL:
OBJECTIVE:

ACTIVITY	PERSON RESPONSIBLE	OTHERS INVOLVED	RESULTS/MEASURE	DATES BEGIN	DATES END
Complete evaluation proposal/plans.	1. N. Woods	1. R. Oxhandler, M.Neill	1. Plans completed.	1. 7/28	9/29
Form task group.	2. N. Woods	2. M.Neill R. Oxhandler	2. Task group convened.	2. By	9/22
Review plans with task group.	3. N. Woods	3. Task Group	3. Plans reviewed/ revised.	3. By	9/30
Devise/approve instruments for: a. student survey b. pre & post testing in FDC 100 c. Future Shoppe data collection d. Advisory Committee activities questionnaires e. dissemination target group questionnaire f. check list for evaluator observation of project activities	4. N. Woods	4. Task Group, Graduate	4. Instruments approved by Task Group.	4. 9/29	10/24
On site interviews with students after admissions counseling.	5. N. Woods	5. Graduate intern	5. At least 30 interviews completed.	5. 10/13	10/24

12

50 =

Nancy Woods
KVCC

EVALUATION WORKPLAN

Instrument	Instrument Status	Evaluation Questions Addressed	Administration Schedule	Administrators	Respondents	Sample	Data Analysis Procedure	Report Available
1. Practicum Preference Form	Completed	B1	before practicum placement	self administered	Student	All students	Frequency, compare to site profile	4 weeks after practicum placement
2. Student Practicum Log Analysis Form	Completed	B1, B2, B3. E1	monthly	self administered	Evaluator	Logs from each student	Frequency of activities, mean time per activity area	End of semester
3. Student Self-Rating Form	Completed	C1, C2, C3, C4	weekly	self administered	Student	All students	Frequency trend over time	Monthly
4. Supervisor Rating Form	Completed	C1, C2, C3, C4	bi-weekly	self administered	Supervisor	All supervisors	Frequency trend over time	End of semester
5. Two step Graduate Questionnaire:	Completed	D1, D2, D3	annually	self administered	Graduates	All graduates	Frequency analysis, central tendency trend by years	End of year
6. Graduate Follow-up Questionnaire: Employers	Completed	D3	annually	self administered	Graduate employers	All; contingent on graduates'	Frequency analysis, central tendency trend	End of year
7. Site Profile Analysis Form	Completed	B1	end of semester	evaluator	Evaluator	One per site	Content analysis; compare to Student Preference	End of semester
8. Clinical Service Review Sheet	Completed	E1	annual	evaluator	Evaluator	Systematic Sample of Records	Frequency central tendency	End of year
9. Contact Record Analysis Form	Ready for field test	A1, A2, A3	end of semester	evaluator	Evaluator	Record for all students	Content analysis frequency, mean time	End of semester
10. Resource Center Usage Form	Completed	A5	end of semester	evaluator	Evaluator	New Materials	Frequency by materials categories	End of semester
11. Drop Out Record Checklist	Completed	C4	annual	evaluator	Evaluator	All drop-out records	Content analysis	End of year
12. Module Feedback Form	Blueprint completed	A2, A3, A5	upon module completion	self administered	Students	All Students	Mean ratings	End of semester
13. Module Test Form	To be developed	A4, B1	upon module completion	self administered	Students	All Students	Mean ratings	End of semester
14. Diagnosis Analysis Checklist	Completed	C1, C2	end of practicum	evaluator	Evaluator	One per student	Content analysis frequency and rating scores	End of semester
15. Diagnosis Observation Sheet	To be developed	C1, C4	end of practicum	faculty	Faculty	One per student	Frequency by client type	End of semester
16. Lesson plan Analysis Sheet	Completed	C2, C3	end of practicum	evaluator	Evaluator	Random: 3 per student	Content analysis frequency, rating scores.	End of semester
17. Classroom Observation Record	Blueprint Completed	C2, C3, C4	end of semester	faculty	Faculty	One per student	Frequency by client type	End of semester

Content —
will the instrument record data of **sufficient** but not excessive specificity and quantity?

Instrument: _____

Evaluation Questions addressed:

Dimensions	Further Definition (if needed)	Items providing information about this Dimension

Dimensions for which additional items are needed: _____

Items to be deleted (those which do not provide any relevant information): _____

Additional Notes: _____

26 (in)

Instrument Analysis Blueprint

Introduction –
is the purpose of the instrument clear? respondents' confidentiality assured? timing and method of returning the instrument specified?

Directions –
are instructions clear and complete? special modes of responding explained? relevant materials for the task specified?

Organization –
are items grouped in a logical order (e. g., by content, type, etc.)?

Format –
are items spaced to facilitate reading and accurate responding?

EVALUATION DESIGN

Evaluation Concern	Evaluation Question	Design Referent	Why the question is important

EVALUATION DESIGN

Evaluation Concern	Evaluation Question	Design Referent	Why the question is important

-2AK-

Procedure: Observe and record behaviors of others.

Description: An observer (a participant or nonparticipant) records the behavior of someone in a particular setting. Specific behaviors are categorized or counted.

Remarks: This is the most direct measure of behavior. An experienced, trained or perceptive observer can pick up subtle occurrences or interactions. A well constructed observation instrument standardizes the observations and can be objectively scored. Problems with the mode are that it is open to observer inconsistencies, there is a possibility of an inadequate sampling of observed events, and the observer's presence may cause an artificial situation. The instrument itself may overlook meaningful behaviors and lack flexibility by lumping unlike behaviors together. Behaviors listed must be observable. Usually, observers must be trained in order to reliably observe and categorize behaviors.

Example Instrument Excerpts:

1. Using the data sheet below, record whether each of the behaviors listed and defined here occur during the time intervals specified. Place a mark in the appropriate cell each time you observe one of the three listed behaviors.

   O = Out of Seat (David) *Definition:*

   R = Reprimands (by Teacher) *Definition:*

   N = Negative Comments (Mary) *Definition:*

	12:15	12:30	12:45	1:00	1:15	1:30	1:45	2:00	2:15	2:30	2:45	3:00
O												
R												
N												

2. Did the tester greet the pupil at the beginning of the session?

   ( ) yes ( ) no

3. Circle below each phrase spoken by the tester during the feedback session.

   - "wrong"
   - "made a mistake"
   - "incorrect"
   - "failed"

-JAL-

Procedure: Record own behaviors.

Description: The respondent maintains a record of events or behaviors that have occurred to himself or another on a regular (daily, weekly, etc.) basis.

Remarks: Logs can provide a very comprehensive record of process since they allow data to be collected from different situations. Like other self-reports, the results are dependent on the respondent's security and honesty.

In constructing logs, care must be taken to construct precise and well-defined categories. If the respondent does not know in what category to record certain events or behaviors, the log quickly becomes inaccurate. For this reason, training of respondents and field testing are important.

Logs can be time-consuming to complete. They may also be difficult and time-consuming to analyze. Without proper structure, there may be a variation in level of detail provided by the respondent.

Example Instrument Excerpts:

Record the number of hours spent this week in each activity listed below:

Activity	Time Spent *(whole hours only)*
A. Discussion with practicum cooperating teacher	_____ hrs.
B. Non-Instructional involvement with pupils	_____ hrs.
C. Preparation for instruction	_____ hrs.
D. Instruction	_____ hrs.
E. Paperwork	_____ hrs.
F. Conferences	_____ hrs.

25c (1n)

Procedure: Conduct a survey.

Description: The respondent is asked to record or categorize self-characteristics and behaviors or those of others, or of events and characteristics of his environment.

Remarks: Surveys are a relatively inexpensive means of collecting alot of information in an easily assimilable and analyzable form. They are quite easy to administer and can reach a large sample of respondents. Return of completed instruments may be a problem.

Since the survey has little control over the response situation, great care must be taken to write well-defined and explicit items and directions. Respondents must be told exactly how to respond; when answers are open-ended (supply-type items), the respondent must be told in what *units* to respond (days, hours, etc.). When the respondent is asked to report frequencies of events or behaviors, items must be written explicitly enough to ensure that the respondent can discriminate an instance from a non-instance for whatever event he is asked to report on.

Respondents are generally cooperative, but instrument length will affect return rate, as will instrument clarity, readability and format. Accuracy of reported data is open to question, so care must be taken to ensure items will elicit honest, unbiased responses.

Example Instrument Excerpts:

1. How frequently did your supervisor visit you to discuss your practicum progress?

   _____ days per week

2. Which of the below is *most* true of your supervisor's visiting schedule?

   ( ) visits almost daily
   ( ) visits at least once each week
   ( ) visits a few times a month, but not every week
   ( ) rarely visits at all

3. Of the topics listed below:

   a. Place an (*X*) in front of those definitely *not* needed for your work.
   b. Place a (✓) in front of those definitely important to your work.
   c. Place a (*O*) in front of those which are inconsequential to your work.

   Leave *blank* any topics about which you're not sure.

- JAN -

Procedure: Administer objective tests.

Description: Respondent selects or generates responses to given questions intended to assess knowledge or understanding of cognitive variables.

Remarks: Knowledge tests is probably the area in which we have the most expertise in controlling for error and thus increasing the consistency of the measure. Validity is often accounted for by the use of a content "blue-print" or table of specifications. There are two major types of tests:

Supplied Answer:	Essay
	Completion
	Short Response
	Problem Solving
Selected Answer:	Multiple Choice
	True-False
	Matching
	Rank Ordering

For selected answer tests, most of the time is spent in the construction. For supplied answer tests, the most time is spent on achieving objectivity in the scoring. Item analysis techniques (for multiple choice items) enable increasing reliability, and, when used with item pools, allow for constructing tests of varying difficulty levels.

Example Instrument Excerpts:

1. Which of the methods listed below is most effective with a Schwartz-syndrome adult? (check one)

   ( ) Houzon-Furst questions
   ( ) Plumb's probing
   ( ) Medley's method
   ( ) Covert's operation

2. Name three (3) reasons why behavior modification techniques may prove ineffective with the Schwartz-syndrome adult.

   a. _____

   b. _____

   c. _____

Procedure: Collect ratings: self or other.

Description: The respondent observes or reflects on some event, person or product and records or categorizes judgments about the quality or extent of characteristics.

Remarks: Unless care is taken in the construction of rating scales, they often tell more about the respondent than about what is being rated. When dealing with past events, the technique is dependent also on the accuracy of the respondent's memory. For self-rating scales, the accuracy of the results can depend on the respondent's security, self-awareness, and honesty. Rating scales, since their variables do not translate into behaviors, may not provide useful feedback to a person being rated. Finally, rating scales are notoriously susceptible to "halo effect" in which an initial general impression will affect all variables being rated.

Rating scales are easy to construct, and are usually easy to complete. Respondents are generally cooperative. The scales are easily scored and data are easily tabulated.

Example Instrument Excerpts:

1. How are others affected by his appearance and manner?
   - ( ) sought by others
   - ( ) well liked by others
   - ( ) liked by others
   - ( ) tolerated by others
   - ( ) avoided by others
   - ( ) no opportunity to observe

2. When an answer is wrong, the student _____ tells the class immediately. (check one)

   ( ) almost never ( ) seldom ( ) sometimes ( ) often ( ) almost always

3. Please rate this person on the following scale: 1 = excellent; 2 = very good; 3 = adequate; 4 = somewhat inadequate; 5 = inadequate. Circle N/A if the item is not applicable to this person or position.

a. Skill in diagnosis	1	2	3	4	5	N/A
b. Thoroughness in keeping records	1	2	3	4	5	N/A
c. Ability to relate with other staff	1	2	3	4	5	N/A
d. Skill at presenting cases	1	2	3	4	5	N/A

- JAP -

Procedure: Analyze products.

Description: The respondent performs a directed analysis of some product (usually a written document) then rates or records certain characteristics of that product as directed by the product-analysis instrument items.

Remarks: This procedure is very similar to behavioral observation, except the object of measurement is a particular document or product. With a well constructed instrument, it can yield very objective and reliable data about the product being measured. Much care must be taken to ensure that the characteristics to be observed and rated or recorded are carefully defined, so that the respondent can make reliable categorizations. Often, training of respondents is required. As with most procedures, the more time spent in instrument development, the easier and quicker will be the analysis. Inter-rater reliability may be a problem, as can intra-rater consistency (such as may occur when content analysis is tedious and a rater has many products to analyze and rate).

Example Instrument Excerpts:

1. Does the lesson plan contain an evaluation procedure?

   ( ) yes ( ) no

2. Check from the list below those behaviors that are cited in the student's report as supporting a Schwartz-syndrome diagnosis:

   ( ) leaping in and out of chairs
   ( ) making insulting comments to small dogs
   ( ) overtightening jar lids
   ( ) pretending to be asleep when spoken to
   ( ) inhalation of fumes from burning vegetal matter

3. What proportion of behaviors cited are accompanied by explanatory comments?

   ( ) all ( ) few
   ( ) most ( ) none

Procedure: Analyze existing records.

Description: Existing records are accessed or collected, then analyzed and their relevant aspects recorded on a form designed for that purpose.

Remarks: The environment in which our education programs operate often contains all sorts of records, files, etc., which could be assessed for measurement purposes. This procedure is quite unobtrusive and objective; further, the measuring of records interacts minimally with the records themselves—the records, files, etc., often "tell it like it is." A problem can be that records may not be accurately maintained or complete—thus they may *not* an accurate measure of what you're after. If they can be trusted, however, and they seem to be a powerful indicator of what you wish to know, it can be a relatively cheap and easy process to measure them.

Example Instrument Excerpts:

1. Looking at the routing slips on each of the five (5) IEP's you selected:
   a. How many are signed by the parent? _____ number signed
   b. How many are dated prior to Jan. 1? _____ number dated

2. For the student teacher's class record:
   a. How many appointments with the master teacher are recorded for each month below?

MONTH	NO. OF APPOINTMENTS
December	_____
April	_____
June	_____

3. How often has the teacher signed out the film, "Cattle-probes in the Classroom"?
   _____ no. of times

Procedure: Produce an inventory.

Description: Respondent counts, measures or categorizes certain tangible objects, recording results on a checklist or other form.

Remarks: The tangible objects checklist is easily administered and since it deals with tangible objects, there is usually very little question about error or judgmental problems entering the picture. The data obtained is easily analyzed typically in frequency counts. Care must be taken to ensure that the respondent can in fact reliably observe and categorize given the items listed.

Example Instrument Excerpts:

1. Check those items which you see in evidence in the classroom:

( ) projector screen	( ) opaque projector
( ) record player	( ) work benches
( ) tape recorder	( ) rugs
( ) filmstrip projector	( ) basal texts
( ) overhead projector	( ) flying saucers
( ) dead dogs	( ) library materials

2. Record in the space below the number of mathematic books on the student's bookshelf:

_____ books

3. How many persons were present?

( ) 3 or less

( ) 4-10

( ) 11 or more

The culminating activity for each participant was to develop an evaluation plan for an in-progress or projected program for young children in his/her school or public library.

A short conference evaluation was conducted at the conclusion of the final session. In general, interest was strong and some excitement about the process and procedures of evaluation was generated among those who attended. The responses are summarized as follows:

SUMMARY OF CONFERENCE EVALUATION

Working Conference on Measurement and Evaluation
Western Michigan University
April 17-18, 1979

The objectives of this conference are stated below. Please indicate the degree to which these objectives have been met for you.

A. To improve skills in program evaluation for practicing professionals.

very high	high	average	low	very low
3	7			

B. To identify steps involved in the evaluation process.

very high	high	average	low	very low
5	5			

C. To develop a tentative evaluation plan for one in-progress or projected program.

very high	high	average	low	very low
2	5	3		

D. Please rate the following features or aspects of the working conference.

	very good	good	fair	poor	very poor
Arrangements					
1. Meeting room	6	4			
2. Meals	7	2			
3. Lodging	5	1			
4. Parking	6	4			
Consultants and Resource Personnel					
1. Style used	9	1			
2. Information given	7	3			
3. Accessibility	7	3			
4. Interaction among participants	7	3			

E. What features of the conference caused you to attend. Please check as many as apply and add comments:

Topic, content	9	Incentive (mileage, reimbursement housing sub pay)	6
Location	5		
Timing (April, weekday)	4		

COMMENTS:

We had to bring a program that was already plannned... feel we were somewhat prepared.

Informative.

Interesting.

We really needed this.

I couldn't have come if you hadn't paid for the substitute.

Was interested in developing skill in program planning and evaluation.

F. Please rate the overall usefulness of this workshop to you.

very high	high	average	low	very low
3	7			

G. Any general comments, observations? Please let me know.

I didn't want to come at first but was glad I did.

Would have liked to have gotten into the actual writing of objectives.

After evaluation, how does one report the findings? Is there any formal or detailed way?

Thank-you.

We could have used several more hours of work and discussion

Very enjoyable.

Am anxious to apply it in my own work.

It should be valuable in my programming of children's activities.

Should help us demonstrate to the public that services we provide are valid and are serving to fulfill some of the needs in the community.

Evaluation of the workshop should be given in 6 months - what have the participants used to improve library services to children.

The workshop leaders were super - a very talented group.

Variety, approach and expertise of the consultants and resource personnel was indicative of the high quality of the working conference.

Overwhelmed.

Real impact is on all programming - our total program effort.

We must get our act together and hopefully this workshop will do the trick.

Conference Follow-Up -

Because the conference was convened later than had been projected in the proposal, a no-cost extension was requested (and granted) so that participants could have time to implement the program evaluation developed within the conference session. Contact was maintained during the interval. (See Appendix D for an example.) The gasoline emergency mecessitated a minor departure from the budgeted plan for site visits. Feed-back was accomplished by means of a mail/telephone mechanism. A form was devised whereby each participant could report in terms of the aspects of evaluation which were stressed during the working conference:

Identify and describe briefly the project, program. Include assessment of need and I/P/O.

INPUT(S) (Resources)	PROCESS (Method)	OUTPUT(S) (Outcomes, goals that were met)

WHAT - is it that you wanted to know as a result of this program?

WHY - do you need to know?

HOW - was the evaluation accomplished --- techniques, tools, method?

WHO - will need, who will be using the information --- users, dissemination?

PRODUCT - what is the format, style of the completed evaluation?

____ Copy for Participant ____ Copy for Director

Programs reported on span a wide range of traditional and non-conventional styles:

-Promoting within the school a public library summer reading program

- Expanding publicity program

-Effectiveness of summer reading clubs

-"Cool Cooking" to reach children who are not library users

-Magic shows as a means of reaching a specified age group

-Effectiveness of a central service center to support branch library activities

Techniques of measurement and evaluation included:

-gathering statistics

-interviews

-graphs of measurable activities

-written qualitative reports

-"before/after" data

-age/sex profile of attendees

A telephone interview at a pre-determined time was conducted with each participant. The same five questions were directed to each person:

(1) Was evaluation plan built in at the time of planning the program?

(2) Did the evaluation process alter anything during the implementation of the program?

(3) What would you do differently another time?

(4) What questions will the outcomes of this program evaluation answer?

(5) Was there anything you did specifically because of what you learned at the conference?

Without exception participants reported that decisions about evaluation procedures were included in the original planning for their respective programs. There were two instances of formative evaluation which resulted in departures from the original plan of action. In response to the query about what might be done differently another time, the replies ranged from simplistic concerns such as changing the calendar dates for a program offering to getting input from staff earlier in the planning phase. Outcomes of program evaluation will be used to: (1) justify additional staff, (2) solidify cooperation between school and public library, (3) establish need for transportation, and (4) verify a pilot program for use in a larger unit. Responding to what occurred specifically as a result of what transpired at the conference, participants indicated that they had used some of the planning instruments introduced at the conference; forms were used with their staffs with good results. One person reported that a half day workshop on evaluation will be implemented with personnel in a regional system. There was general indication of heightened awareness of the need for and the techniques of evaluation.

ANALYSIS AND CONCLUSIONS

This research and demonstration project testifies to the need for greater awareness of both purpose and process of evaluation for library practitioners who work with children in school and public libraries. There is also an

identified need for non-administrative personnel to be more involved at the early phases of planning programs. Among the participants in this project and among those who were contacted as potential participants, there was rather general protest about interest on the part of administrators being "turned on and off". Little interest is evidenced about positive results; concern for evaluation surfaces in times of economic or professional stress.

Since there was a wide spread re: level of professional sophistication and experience represented among the participants from sub-assistant to branch professional to system (regional) coordinator, there was reflected an equally wide range of perception about the role and nature of the process of evaluation. The constructive aspects of evaluation were not, prior to this conference, a working mechanism for many of these practicing professionals. Within the scope of the two day session they seemed to be concerned with professional security (i.e. employment) and program justification. There was little evidence that they direct attention to the impact of programming on young children. This is not to say there was apathy, but rather to emphasize the need which this project addressed. The excitement and earnest professional concerns generated were heartening. Feedback after a four month interval revealed no less interest, but some disappointment was expressed about the difficulty of implementing sound evaluation procedures with colleagues "back home" who had not shared the conference experience. The philosophical concept of evaluation as a way of looking at outcomes as they relate to programming goals and objectives, as a way of improving programs by building evaluation measures in at the planning phase instead of "after the fact" is not a universal modus operandi for practicing library professionals. These needs should continue to be included or expanded in library education and in-service or continuing education activities.

SUMMARY AND OUTCOMES

The experience of implementing the research aspect of this project has been invaluable for the Project Director insofar as 1) substantiating the conviction that many practicing professionals need more skills and techniques of evaluation, and 2) confirming the need to draw upon related disciplines for concepts and practices which can be replicated in librarianhsip.

The demonstration aspect fulfilled the projected activities of the proposal with the previously reported changes in calendar and staffing. The difficulty encountered in obtaining a full roster of participants raises the issue of whether an invitational session might have been preferable. The revisions of the projected plan of action due to illness and weather confirms the wisdom of affording a Project Director some latitude in decision making within the duration of a project. In this instance the need to replace two consultants resulted in securing the expertise of four very knowledgeable persons who complemented each other with highly productive results. Their ability to relate professionally and personally to each other and to participants was clearly a plus factor. They were a substantial replacement for the consultants originally cited in the proposal.

In summary, the projected activities of the proposal have been accomplished. The survey of the literature on evaluation in several related fields has resulted in a publishable document. Negotiation is underway to publish this in its present format or in an expanded version. The impact

of the working conference is limited to those who participated, but hopefully the professionals and the programs in their respective libraries will be influenced as well. Contact will be maintained informally by the Project Director so that any subsequent outcomes can be reported. Inclusion of the findings of this research and demonstration project will most certainly find their way into the course content of Western Michigan University School of Librarianship offerings. Raising the level of consciousness re: evaluation needs and techniques is without doubt a prominent outcome of the project.

APPENDIX A:

"Excerpts from the Authorities: a Fact Sheet..."

EXCERPTS FROM THE AUTHORITIES:* A FACT-SHEET
Prepared for Measurement and Evaluation Conference Participants
Western Michigan University
April 17-18, 1979

DEFINITIONS RELATING TO EVALUATION

. . . goal effectiveness = "overall assessment of program impact and effectiveness with emphasis on extent to which programs are successful in achieving basic objectives." (19)

. . . process of assessment is designed to provide information about past and present operations and effectiveness, in order to assist in making decisions about the future (16)

. . . types of programs:

-demonstration = experimental
-operational = continuing (19)

. . . demonstration programs develop model or prototype; test on a small scale (24)

. . . operational programs attempt to meet some existing need by whatever means can be realistically established; requires feedback and constant revision (24)

. . . difference between goals and objectives: objectives are operational definitions of goals; therefore they can be measured (10)

. . . formative evaluation produces information that is fed back during the program during the development in order to improve it; summative evaluation provides information about effectiveness to decision makers (25)

. . . measurements are indicators of outcomes indicators are developed to measure the extent to which goals are achieved (25)

. . . cost effectiveness = optimal output for a given input; program effectiveness = careful assessment of goals (10)

. . . program includes objectives, activities, resources.
statement of program consists of:
resources a, b, c, used to perform activities d, e, f, which in turn are designed to attain objectives g, h. (6)

. . . objective includes:
what - the nature of situation or condition
extent - quantity or amount to be attained
who - particular persons or group for which attainment is desired (6)

. . . evaluative criteria = basis on which decisions are made about means towards ends (24)

. . . "rolling evaluation" = measure of one objective, and then continue with others (19)

*Numbers in ( ) refer to titles cited in the bibliography.

CHARACTERISTICS OF EVALUATION

. . . measures effects (outcome, not efficiency) of a program against the goals it set up to accomplish (25)

. . . a means of contributing to subsequent decision making about the program and improving future programming (25)

. . . seeks change while the natural tendency of program operation is to resist change (24)

. . . judges merit; examines and weighs a phenomenon against some explicit or implicit yardstick; establishes clear and specific criteria for success (25)

. . . whoever wants to evaluate has some reason in order to make some kind of decision or to serve some personal or group purpose (20)

. . . two classes of program evaluation:

Operational	Impact
How well did the program perform?	Are the observed performances attributable to the program or its components?

. . . many aspects of evaluation hinge upon individual and group motivations:

- -may govern the kind of program development
- -may govern the kinds of assumptions to be made (or rejected)
- -may govern the standards or judgment against which effectiveness will be measured
- -may govern the level of analysis and what data to include
- -may govern who shall be involved, the type and manner of controls, the nature and the frequency of feedback and reporting (20)

. . . value of evaluation depends upon quality of the information expected---how well defined and the specific nature of the data (5)

. . . program evaluation directly challenges the worthwhileness of specific activities or performance of specific individuals; the focus is on success or failure, therefore staff is highly anxious (24)

. . . Operational evaluation is a comparison process during which the performance of a program is judged; based on:

- -program's attained results
- -base to which measured results can be obtained

. . . absence of a model results in being mired in status quo or in blindly pursuing a course before merits are assessed (10)

. . . is evaluation always good?
do the reasons always justify the effort involved?
"evaluation is like salt--nearly all foods taste better with some salt. Some foods need more salt than others, but too much salt can ruin the meal." (20)

WHY EVALUATE?

. . . to affect resource allocation
to improve strategies (16)

. . . to communicate to others what the program is like; evaluation will present diversity, complexity, successes, failures, plus differing value perspectives

. . . to provide specific information about the program because a decision about its future is necessary

. . . to provide a means by which feedback about the intent of the program, its activities and its outcomes are provided (24)

. . . may be used:

-to monitor a steady state or determine a need for correction
-to identify alternatives in a problem situation
-to weigh alternatives in terms of relative gains or losses
-to determine corrective action and risks involved in change (24)

. . . "the why" helps a person identify:

-kind of evidence needed
-amount and quality of evidence required
-methods of measurement to be used (20)

. . . the impact of any program, how small or modest can be constructively evaluated (5)

. . . evaluation = process of determining the value or worth of something relative to a given purpose or standard; process includes:

-making decisions
-drawing together evidence
-weighing pro's and cons
-selecting courses of action (20)

WHAT IS BEING EVALUATED?

. . . three key elements must be present before evaluation can take place:

-goal and objective(s) which are considered to be desirable or has some positive value
-planned program of deliberate intervention which hypothesizes that the program is capable of achieving goal and objective(s)
-method for determining the degree to which goal or objective(s) have been attained as a result of the planned program (24)

. . . differences among programs: scope, size, duration, clarity and specificity of program elements (25)

. . . certain quantitative aspects can be easily measured:

-how many children
-how many parents
-who is affected

but you are also evaluating what actually happens to values, attitudes, understandings of children, parents, community at large (21)

. . . reasons for evaluation are either organization-oriented or personally-oriented

Organization-oriented reasons:

-to demonstrate that program is worthwhile
-to determine if program is moving in right direction
-to determine whether original need(s) for program are being justified
-to justify past or projected expenditures
-to demonstrate cost in terms of money and/or human effort
-to demonstrate effectiveness of a program
-to gain support for program expansion
-to compare different kinds of program approaches or method in terms of effect (20)

Personally-oriented reasons:

-evaluation is the "thing" to do", i.e., status, acceptance of peers and superiors
-may wish to evaluate to bring favorable attention, better budgets or better staff
-a way of making job easier or more interesting
-a necessary step for promotion or merit
-a vague or urgent need to see progress (20)

HOW DO WE EVALUATE?

. . . steps in measuring program effectiveness:

-describe the program
    -specify objectives
    -specify activities
-measurement
    -kinds of evidence needed
    -when to measure (short-range or continuous)
    -how to measure, avoid biasing outcomes
-determine effectiveness
    -compare the accomplishment attributable to the program with the accomplishment intended for the program
    -assess the extent to which achievement of the objective can be attributed to activities performed in the program
-use of findings
    -locate program difficulties by measuring variables (resources, activities, objectives
    -reasons for less than effective programs:
        -resources not used as planned
        -assumptions (needs) not valid
        -activities not performed as planned (6)

. . . choice among evaluation procedures depends on a balance of:

-detailed objectives
-range and nature of possible outcomes
-how outcomes will be measured
-cost of evaluation procedure (5)

. . . collect data by: asking people, observations, program records

. . . types of measures:

-effects on persons served, changes in program participants
-measuring effect on larger system
-measuring effect on the public (community) (25)

. . . components of evaluation:

-identify purposes
-audience
-relevant issues
-resources available (11)

. . . kinds of evaluative questions:

Appropriateness

importance of problems selected and relative emphasis or priority

Adequacy

objectives are directed toward eliminating the problem(s) which gave rise to the program

adequacy of program objectives refers to how much of the problem the program is supposed to eliminate

scope of objectives may need to be reduced to satisfy constraints

Effectiveness

extent to which pre-established objectives are satisfied

Efficiency

cost of resources expended in attaining objectives when attainment is to be accomplished

Activity

includes work or occasion in the attainment of objective; diverse efforts

activity is "flapping wings"; objective is being at a desired place

activities consume program time and resources; objectives to not

program planning parallels cause/effect relationship: a given set of activities will lead to attainment of objectives

evaluation tests the hypotheses, "if. . .then"

Resources

include personnel, funds, materials, facilities available (or projected) (6)

. . . steps in evaluation:

-find out the program's goals
-translate the goals into measurable indicators of achievement
-collect data on the indicators
-compare data with goal criteria (25)

. . . evaluator must decide which goals to evaluate:

-usability and practicality--potential for utilization
-relative importance--make priorities
-incompatibilities
-short-term or long-term goals
-yardsticks
-unanticipated consequences (25)

. . . evaluation design worksheet (sample):

Purpose

why evaluate

Audience

who is evaluation for

Issues

what questions should the evaluation address

Resources

what are available for the evaluation

Evidence

what evidence should be collected

Data Gathering

how is evidence to be collected

Analysis

how is evidence analyzed

Reporting

how can findings be reported (11)

USERS OF EVALUATION

. . . users include:

-funding organization
-service staff
-patrons (25)

. . . various kinds of people are interest; respond to all

-participators
-funders
-administrators
-community leaders (11)

. . . program directors want to know what strategies were successful, which are achieving results most efficiently what features of the program are successful (25)

. . . follow-up of evaluation should be in the hands of program staff; evaluator presents findings but does not enforce change

. . . findings are much more likely to be used if program staff is given part in the evaluation and direct responsibility for implementing the findings

. . . stress program improvement instead of pass/fail judgment (24)

. . . attempt to improve based on feedback by:

-allocating resources differently
-changing procedures
-reorganizing (18)

. . . increase potential use of evaluation findings:

-specify links from program input to outcome
-analyze effectiveness of program components or alternative approaches rather than all-or-nothing, go/no go assessment of the total program (24)

. . . legitimate by-products of evaluation--desirable secondary benefits:

-clarification and re-examination of objectives and underlying assumptions
-more careful analysis of content and operation of the program to specify what are the essential aspects of the program
-identify and define the target population
-re-evaluate program especially in terms of how principles are translated into practice
-formulate hypothesis for further study to fill gaps which affect program planning
-increase staff morale and interest (24)

WORDS TO THE WISE . . . ADVICE TO EVALUATORS

. . . evaluation often answers questions which no one is asking

. . . think about evaluation needs before conducting the program

. . . evaluations are better if planned early in the planning of of the program (11)

. . . no easy answer to administrative evaluation problems; needs supportive climate and staff dedicated to program improvement (24)

. . . evaluation is best when planned for and monitored; program personnel's commitment to self examination and improvement are musts (11)

. . . evaluator should not rely only on his/her own powers of observation, judgment, and responding; should enlist aid and make repeated observations (21)

. . . how do we maintain flexibility within structure?

-planner thoroughly commits himself to decisions which plan dictates
-operationally oriented manager tries to overcome difficulties as they occur (10)

. . . be definite in terms of time and personnel commitment (19)

. . . find satisfaction in planning while recognizing the limitations of working in a world of flux; if you "roll with time, you will find neither business as usual (blind quest for efficiency) nor change for the sake of change (thoughtless faddism) (10)

. . . mis-uses of evaluation:

-eye-wash---attempt to justify a weak program by selecting only those aspects that "look good"
-white wash--cover up failure or errors by avoiding objective evaluation--"vindication replaces verification"
-submarine--attempt to torpedo or destroy a program regardless of effectiveness--"politics replaces science"
-posture--a gesture of objectivity--"ritual replaces research"
-postponement--delay needed action by pretending to seek facts--"research replaces service" (24)

BIBLIOGRAPHY

1. Altman, Ellen O. "Measurement and Evaluation." in American Library Association. ALA Yearbook, 1977, pp. 201-2.

2. Aronson, Sidney H. and Clarence, C. Sherwood. "Researcher versus Practitioner: Problems in Social Action Research." Social Work, v.12, no. 4, 1967. pp. 89-96.

3. "Assumptions Underlying Evaluation Models." Educational Researcher, v.7 no. 3, March 1978. pp. 4-12.

4. Blalock, Hubert M. Jr. An Introduction to Social Research. Englewood Cliffs, Prentice, 1970.

5. Davidson, Park O. and others. Evaluation of Behavioral Programs. Research Press, 1974.

6. Deniston, O. L. and others. "Evaluation of Program Effectiveness" in Schulberg, Herbert C. and others. Program Evaluation in the Health Sciences. Human Sciences Press, 1977. Pp. 219-242.

7. DeProspo, Ernest R. "Measurement and Evaluation." in American Library Association. ALA Yearbook, 1976, pp. 231-2.

8. "Evaluation and Analysis: Who Needs It? You Do!" PLA Newsletter 16:16-17, Fall, 1977.

9. "Evaluation Research: A Bibliographic Overview." Planning Librarians Exchange Bibliography, #975, February, 1976.

10. Goldberg, Robert L. A Systems Approach to Library Program Development. Scarecrow Press, 1976.

11. Grotelueschen, A. D. and others. An Evaluation Planner. Urbana, University of Illinois, 1976.

12. Herzog, Elizabeth. Some Guidelines for Evaluative Research. Washington, D.C., Department of Health, Education, and Welfare, 1959.

13. Jones, Richard R. "Design and Analysis Problems in Program Evaluation", in Davidson, P. O. and others, eds. Evaluation of Behavioral Programs. Champaign, Research Press, 1974. Pp. 1-32.

14. Knutson, Andie L. "Evaluation for What?" in Schulberg, Herbert C. and others. Program Evaluation in the Health Fields. Human Sciences Press, 1977. Pp. 42-58.

15. Lancaster, F. W. and Jonich, M. J. Measurement and Evaluation of Library Services. Information Resources Press, 1977.

16. Levine, Robert A. and Williams, A. P., Jr. Making Evaluation Effective: a Guide, a Report prepared for the Department of Health, Education and Welfare, #R-788-HEW/CMU. The Rand Corporation, May, 1971.

17. Massey, Julie and Koziol, Stephen Jr. "Research on Creative Dramatics." English Journal, February, 1978, p. 92.

18. Raizen, Senta A. Evaluating NIE The Rand Corporation, February, 1973.

19. Roberto, E. L. Strategic Decision Making in a Social Program. Heath, 1975.

20. Schulberg, Herbert C. and others. Program Evaluation in the Health Fields. Human Sciences Press, 1977.

21. Stake, Robert. Evaluating the Arts in Education. Columbus, Ohio, Merrill, 1975.

22. Stufflebeam, Daniel S. "The Use and Abuse of Evaluation in Title III." Theory into Practice, June, 1967, pp. 126-33.

23. Stecher, G. "Library Evaluation: a Brief Survey of Studies in Quantification," Australian Academic and Research Libraries 6:1-19, March, 1975.

24. Suchman, Edward A. "Action for What? A Critique of Evaluation Research". Chapter 5 in O'Toole, R. Organization, Management, and Tactics of Social Research. Schenkman Publishing Co., 1971.

25. Weiss, Carol. Evaluation Research, Englewood Cliffs, Prentice, 1972.

26. Winnick, Pauline. "Evaluation of Public Library Services to Children." Library Trends 22:361-76, January, 1974.

Compiled by:
Marguerite Baechtold
Project Director
Early Childhood Library
Programming

APPENDIX B:

Revised list of Staff for Working Conference

WORKING CONFERENCE ON MEASUREMENT AND EVALUATION

OF

EARLY CHILDHOOD LIBRARY PROGRAMMING

CONVENOR: Marguerite Baechtold

Associate Professor

School of Librarianship

and

Director, Early Childhood
Library Programming Project

CONSULTANTS:

Dr. John Rizzo
Professor of Management
WMU
and
Consultant for Library
Management

Dr. Subhash Sonnad
Associate Professor
of Sociology
WMU

Nancy Woods
Management Studies
Kalamazoo Valley
Community College

RESOURCE PERSONS:

Roger Hartz
Doctoral Candidate
Educational Leadership
WMU

Howard Major
Doctoral Candidate
Educational Leadership
WMU

APPENDIX C:

Announcements, Application and Confirmation Forms,
and List of Participants

EVALUATE! JUSTIFY! MEASURE! PROVE! DEMONSTRATE!

If these are terms which flash in and out of your professional nightmares, there may be help in sight for you!

As part of a federally funded research and demonstration project in Library Programming for Early Childhood, a working conference on measurement and evaluation of such programs will be convened on April 17 and 18, 1979 on Western Michigan University campus. Because you have indicated interest in library programming for young children, you are being invited to apply for inclusion in what promises to be a one-of-a-kind opportunity. If accepted, you will be among twenty selected librarians who will benefit from the expertise of three consultants who have unique competencies in the area of measuring and evaluating programs. Each participant will bring a projected or in-progress program for which evaluation will be developed during the conference.

Enclosed is an application form and a tentative agenda for the conference. If you wish to be considered, please fill out the tear sheet and send it to me by return mail. Participants will be notified by phone or mail so you may promptly make arrangements to attend. Since this is part of a federally funded research and demonstration project, some subsidy can be provided to off-set part or all expenses incurred as follows: mileage, motel, meals, and a substitute to cover your work responsibility if necessary.

For further information, telephone (616) 383-1849 Business
(616) 381-7670 Home

or send the application form without delay to:

Marguerite Baechtold, Director
Early Childhood Library Programming Project
School of Librarianship
Western Michigan University
Kalamazoo, Michigan 49008

T E N T A T I V E A G E N D A

for

WORKING CONFERENCE ON MEASUREMENT AND EVALUATION

of

EARLY CHILDHOOD LIBRARY PROGRAMMING

Sponsored by School of Librarianship, Western Michigan University
under the terms of a grant funded by U.S. Office of Education, HEA Title IV-B

TUESDAY, APRIL 17, 1979		WEDNESDAY, APRIL 18, 1979	
1:30 p.m. to 4:30 p.m.	Convene in WMU Student Center Room 157 Greetings and Introductions Presentations on principles and techniques of evaluation by Conference Consultants: Dr. Ruth Ekstrom, Research Scientist, Educational Testing Service, Princeton, New Jersey Dr. John Rizzo, Professor of Management, WMU and consultant for library management Nancy Woods, Management Studies, Kalamazoo Valley Community College	9:00 a.m.	Convene in WMU Student Center Work Session - - with Consultants - - with other participants
		12:00 p.m.	Adjourn for lunch
		1:00 p.m.	Feedback Session
5:45 p.m.	DINNER	3:30 p.m.	Conference Summary
7:00 p.m. to 9:00 p.m.	Participants identify programs	4:00 p.m.	Conference Adjournment

- - - -

WESTERN MICHIGAN UNIVERSITY
SCHOOL OF LIBRARIANSHIP
KALAMAZOO, MICHIGAN 49001

YOU'RE ON OUR LIST!

This communication will notify you of your acceptance for the Measurement and Evaluation Conference sponsored by the Early Childhood Library Programming Project.

Essential Facts:

1- Tuesday, April 17 1:30 p.m.
We shall convene in Room 158, University Student Center. Plan to reach campus by 1 p.m. to allow time for parking and locating Room 158 in Student Center.

2- Best bet for parking: the parking ramp on W. Michigan Avenue, just east of Student Center. Parking Permit mailed in a separate mailing.

3- Dinner on Tuesday and lunch on Wednesday will be provided in the President's Dining Room, USC.

4- Rooms have been reserved at Red Roof Inn US 131 and Stadium Drive near the campus for those who have requested overnight housing. You may check in before 1:30 session, before dinner, or room will be held for you until the close of the evening session at 9 p.m.

5- Please be prepared to note actual mileage; otherwise "best-guess-estimate" will be used for re-imbursement.

6- Bring with you a written description of a projected or in-progress program to use for your exercise in evaluation.

7- If a substitute is needed to release you from your work responsibility, please have with you the person's name, home address, and social security number so we can arrange for prompt payment.

8- Any questions? ? ?
Call Marguerite Baechtold
Project Director
(616) 383-1849 Office
(616) 381-7670 Home

MB/sb

WESTERN MICHIGAN UNIVERSITY

SCHOOL OF LIBRARIANSHIP

KALAMAZOO, MICHIGAN
49001

April 25, 1979

Dear

I want to thank you for taking time from your professional responsibilities to participate in the Measurement and Evaluation conference, April 17 and 18. The groundwork for some important professional development was laid during the two days. I hope you will find opportunity to build some of those considerations into planning and evaluating your programs on the job.

A secondary reason for writing is to alert you to the fact that the terms of the grant provide for follow-up to the conference. I want, therefore, to let you know that I shall be contacting you again within the next three months to learn about the ways you have been able to apply some of the concepts, techniques, and procedures to evaluating some aspect of your library program. I shall be in touch by phone or mail when I can be more specific about the follow-up.

Again, thank you for your contribution to a most profitable conference. Best wishes as you proceed with your job.

Sincerely,

Marguerite Baechtold
Project Director
Early Childhood Library
Programming

MB/sb

August 16, 1979

Dear

You will recall that in my letter of April 25, I indicated that I would be communicating with you again "within the next three months" to follow up on the outcomes of the Measurement and Evaluation Conference. To my amazement (and perhaps to yours) three months have already elapsed!

My purpose in writing now is to let you know that I am preparing to gather information about the ways you have been able to apply some of the concepts, techniques, and procedures from the conference in evaluating a library program for young children. To facilitate this, I have prepared a preliminary work sheet, and am requesting that you return this to reach me by September 5. Please provide as much information as you can at this time, and relate it wherever possible to ideas engendered at the conference. I plan to confer by phone with each of the participants between September 5 and 14. The information which you send will brief me in advance of the telephone conference, so our conversation can move directly to the specifics of your program evaluation.

Please don't tell me that you "intended to do thus-and-so, but ran out of time, or money, or help..." This has happened to so many post-conference projects! The time we spent in April seemed to be very profitable and productive for all who were involved. I'm hoping this time will be different, and that we shall be able to report to HEW that some practical and usable evaluation practices have come out of this experience.

Thank you for participating in this project.

Sincerely,

Marguerite Baechtold
Director, Early Childhood
Library Programming Project

P A R T I C P A N T S

I N

WORKING CONFERENCE ON MEASUREMENT AND EVALUATION

OF

EARLY CHILDHOOD LIBRARY PROGRAMMING

BURKHAM, MIMI
Washington Square Branch
Children's Librarian
1244 Portage
Kalamazoo, Mich. 49007

DEYOUNG, GAIL O.
Kent County Library
Ass't Children's Services Coord.
775 Ball Avenue NE
Grand Rapids, Mich. 49503

GRANT, BARBARA
Cass County Library
Head of Children's Services
M-62 North
Cassopolis, Mich. 49031

GUNTHORPE, SUSAN
Marshall Public Library
Assistant Director
111 E. Mansion Street
Marshall, Mich. 49068

HALE, VIVIAN
Cass County Library
Ass't Children's Librarian
M-62 North
Cassopolis, Mich. 49031

MACKELLAR, CAROLE
Webster Memorial Library
Children's Librarian
200 Phelps
Decatur, Mich. 49045

MCALLISTER, MARIANNE
Student--WMU
1519 Howland
Kalamazoo, Mich. 49001

MANSHUM, JOANNE S.
Dansville Agricultural School
Children's Librarian
1264 Adams St.
Lansing, Mich. 48819

SMITH, RODGER
Tutt Branch, South Bend Public Library
Branch Librarian
2002 Miami Street
South Bend, Ind. 46613

VAN LIEU, RUTH M.
Branch County Library
Children's Librarian
10 E. Chicago Street
Coldwater, Mich. 49036

WALDO, MARY LOUISE
Kalamazoo Public Library
Children's Librarian
315 South Rose
Kalamazoo, Mich. 49007

APPENDIX D:

Agenda for Working Conference

AGENDA
for
WORKING CONFERENCE ON MEASUREMENT AND EVALUATION
of
EARLY CHILDHOOD LIBRARY PROGRAMMING

Sponsored by the School of Librarianship, Western Michigan University
under the terms of a grant funded by the U. S. Office of Education, HEA Title IV-B

April 17-18, 1979

GREETINGS

For the School . . .

Dr. Jean E. Lowrie, Director

For the Project . . .

Marguerite Baechtold, Project Director

INTRODUCTIONS

Consultants

Resource Persons

Participants

AGENDA
(Each section to begin with discussion, followed by work session)

WHY EVALUATE?

Why do we need to, why do we want to evaluate?

WHAT IS TO BE EVALUATED?

What do we need to know?

HOW TO EVALUATE?

Techniques, tools, method

DISSEMINATION?

Who will need, who will be using the evaluation?

PRODUCT?

Formats, Styles